

CONTENT

Gorban AN, Tyukina TA, Pokidysheva LI, Smirnova EV. Dynamic and thermodynamic models of adaptation. Phys Life Rev 2021; 37: 17–64.

<https://doi.org/10.1016/j.plrev.2021.03.001>

Selvarajoo K. Searching for unifying laws of general adaptation syndrome. Phys Life Rev 2021; 38: 97–99. <https://doi.org/10.1016/j.plrev.2021.04.001>

Giuliani A. The statistical mechanics of life. Phys Life Rev 2021; 38: 100–102.

<https://doi.org/10.1016/j.plrev.2021.04.003>

Liu R, Aihara K, Chen L. Collective fluctuation implies imminent state transition. Phys Life Rev 2021; 38: 103–107. <https://doi.org/10.1016/j.plrev.2021.04.002>

Red'ko VG. Some aspects of adaptation and evolution. Phys Life Rev 2021; 38: 108–110.

<https://doi.org/10.1016/j.plrev.2021.04.004>

Ueltzhöffer K., Da Costa L, Friston KJ. Variational free energy, individual fitness, and population dynamics under acute stress. Phys Life Rev 2021; 38: 111–115.

<https://doi.org/10.1016/j.plrev.2021.04.005>

Franceschi C. Aging, Inflammaging and Adaptation. Phys Life Rev 2021; 38: 107–110.

<https://doi.org/10.1016/j.plrev.2021.07.001>

Bizzarri M, Pontecorvi P. Critical transition across the Waddington landscape as an interpretative model. Phys Life Rev 2021; 38: 115–119.

<https://doi.org/10.1016/j.plrev.2021.05.010>

Nazarenko T, Blyuss O, Whitwell H, Zaikin A. Ensemble of correlation, parenclitic and synolitic graphs as a tool to detect universal changes in complex biological systems. Phys Life Rev 2021; 38: 120–123. <https://doi.org/10.1016/j.plrev.2021.05.009>

Di Paola L, Leitner DM. Network models of biological adaptation at the molecular scale. Phys Life Rev 2021; 38: 124–126. <https://doi.org/10.1016/j.plrev.2021.05.008>

Vasenina E, Kataoka R, Loenneke JP, Buckner SL. Exercise science perspective. Phys Life Rev 2021; 38: 129–131. <https://doi.org/10.1016/j.plrev.2021.05.005>

Zinovyev A. Adaptation through the lens of single-cell multi-omics data. Phys Life Rev 2021; 38: 132–134. <https://doi.org/10.1016/j.plrev.2021.05.004>



Models of adaptation
Review and discussion
2021

PHYSICS of LIFE
reviews

Allahverdyan AE. Energy dissipation and storage in adaptation and homeostasis. Phys Life Rev 2021; 38: 137–139. <https://doi.org/10.1016/j.plrev.2021.05.002>

Gorban AN, Tyukina TA, Pokidysheva LI, Smirnova EV. It is useful to analyze correlation graphs. Phys Life Rev 2021; In Press. <https://doi.org/10.1016/j.plrev.2021.10.002>

Review

Dynamic and thermodynamic models of adaptation

A.N. Gorban^{a,b,*}, T.A. Tyukina^a, L.I. Pokidysheva^c, E.V. Smirnova^c^a *Department of Mathematics, University of Leicester, Leicester, UK*^b *Lobachevsky University, Nizhni Novgorod, Russia*^c *Siberian Federal University, Krasnoyarsk, Russia*

Received 11 March 2021; accepted 11 March 2021

Available online 17 March 2021

Communicated by L. Perlovsky

Abstract

The concept of biological adaptation was closely connected to some mathematical, engineering and physical ideas from the very beginning. Cannon in his “The wisdom of the body” (1932) systematically used the engineering vision of regulation. In 1938, Selye enriched this approach by the notion of adaptation energy. This term causes much debate when one takes it literally, as a physical quantity, i.e. a sort of energy. Selye did not use the language of mathematics systematically, but the formalization of his phenomenological theory in the spirit of thermodynamics was simple and led to verifiable predictions. In 1980s, the dynamics of correlation and variance in systems under adaptation to a load of environmental factors were studied and the universal effect in ensembles of systems under a load of similar factors was discovered: in a crisis, as a rule, even before the onset of obvious symptoms of stress, the correlation increases together with variance (and volatility). During 30 years, this effect has been supported by many observations of groups of humans, mice, trees, grassy plants, and on financial time series. In the last ten years, these results were supplemented by many new experiments, from gene networks in cardiology and oncology to dynamics of depression and clinical psychotherapy. Several systems of models were developed: the thermodynamic-like theory of adaptation of ensembles and several families of models of individual adaptation. Historically, the first group of models was based on Selye’s concept of adaptation energy and used fitness estimates. Two other groups of models are based on the idea of hidden attractor bifurcation and on the advection–diffusion model for distribution of population in the space of physiological attributes. We explore this world of models and experiments, starting with classic works, with particular attention to the results of the last ten years and open questions.

© 2021 Elsevier B.V. All rights reserved.

Keywords: Correlation graph; Network biology; Adaptation energy; Critical transitions; Training; Limiting factor

* Corresponding author at: Department of Mathematics, University of Leicester, Leicester, UK.

E-mail addresses: a.n.gorban@le.ac.uk (A.N. Gorban), tt51@le.ac.uk (T.A. Tyukina), pok50gm@gmail.com (L.I. Pokidysheva), seleval2008@yandex.ru (E.V. Smirnova).

<https://doi.org/10.1016/j.plrev.2021.03.001>

1571-0645/© 2021 Elsevier B.V. All rights reserved.

1. Introduction

1.1. Correlation graph in stress and crisis

Network medicine and biology is undoubtedly a modern vector for the development of biomedical sciences [100]. Integration and use of huge collections of individualized data is not possible without network representation. Modern information technologies, such as semantic zooming, are used to organize and analyze information about networks of biological processes [59]. New experimental technologies, such as single cell omics (transcriptomics, epigenomics, and proteomics), open up opportunities to analyze large samples of single cells and provide access to phenomena, which were invisible when studied by standard methods that average the data across the multiple cells [30]. They pose new challenges to network data analysis [31,93,172].

Some ideas in network biology have a long history. Analysis of graphs of correlations between biological traits was proposed in biostatistics in 1931 [161]. The vertices of these graphs are attributes, and the edges correspond to sufficiently strong correlations (above a certain threshold). The *correlation graphs* were used to define *correlation pleiades* that are clusters of correlated traits. They were used in evolutionary physiology for identification of a modular structure (blocks of connected traits) [6,10,115,161]. Terentjev explained his idea as follows: “Let us take an object as a set of characteristics and consider all conceivable combinations of their correlative relationships. Even superficial observation shows that the uniformity of such correlation of features is a rather rare case. As a rule, the characteristics are grouped into a few ‘societies’ that can be called ‘correlation pleiades’.”

Technically, it is necessary to distinguish between the core and peripheral elements of the correlation pleiades, where small intersections of different pleiades on the periphery are possible. [48, Chapter 2]. (We refer to [180] for a systematic review of clustering.) Later on, it was discovered that the correlation pleiades of physiological parameters may change under the load. For example, the attributes related to the cardiovascular systems and the attributes of the respiratory systems form different pleiades for healthy people at rest but under significant load (measured in cycle ergometer stress testing) the correlations between attributes of these systems increase and their pleiades may join in one cluster [159].

Correlation clusters are used in various fields, such as human metrology [1] (for predicting unknown body measurements and evaluating some soft biometric data), in genomics (for example, for reconstructing RNA-viral quasi-species [9]), and for many other tasks.

Network biology of stress and disease based on correlation graph analysis was discovered in 1987 [68]. A universal effect in ensembles of similar organisms under the load of similar factors was observed. Represent every organism as a vector of biomarkers (physiological attributes of various types: activities of enzymes, data of biochemical blood analysis, activities of various genes, etc.). Under stress, as a rule, even before the appearance of obvious symptoms of the crisis, the correlation between the attributes of organisms in the group increases, and at the same time, the variance also increases. With the development of stress and the approach of a crisis, the connectivity of the correlation network of biomarkers increases. This effect has been repeatedly rediscovered. Several theories have been proposed to explain it. Now it is ‘time to gather stones together’. In this paper we analyze the experimental and theoretical results on the dynamics of correlation graphs under adaptation and stress. Anticipating a detailed analysis, let us formulate the main conclusions obtained as a result of the review:

- The dynamics of the correlation graph has been proven to be a useful marker of adaptation and stress.
- It can be used in development of the ‘early warning’ signals for anticipating of stress and crisis.
- Some additional experimental data are needed to clarify the dynamics of correlations and variability near deaths. The reports are still ambiguous. The hypothesis is that closer to lethal outcome, the correlations between physiological parameters decay, the variance decreases but the relative variation increases.
- Several different theoretical approaches to this dynamics were proposed. They give qualitative explanation of important features but no approach is completely satisfactory and unresolved issues remain.
- A general unified theory of the effect is needed, inheriting the achievements of existing models and solving open questions.

We discussed both successes and failures of modeling and formulate some open questions.

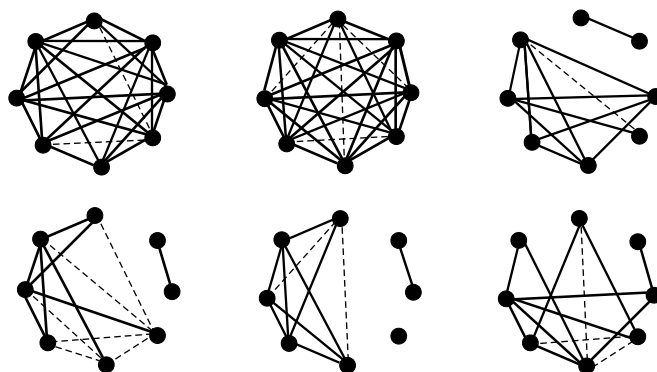


Fig. 1. Correlation graphs of lipid metabolism for newborn babies. Eight lipid fractions were analyzed. Vertices correspond to different fractions of lipids. Edges drawn by solid lines connect vertices with the correlation coefficient between fractions $|r_{ij}| \geq 0.5$. Edges drawn by dashed lines correspond to the correlations $0.5 > |r_{ij}| \geq 0.25$. Upper row – Far North (FN), lower row – the temperate belt of Siberia (TBS). From the left to the right: 1st-3rd days (TBS – 123 and FN – 100 babies), 4th-6th days (TBS – 98 and FN – 99 babies), 7th-10th days (TBS – 35 and FN – 29 babies).

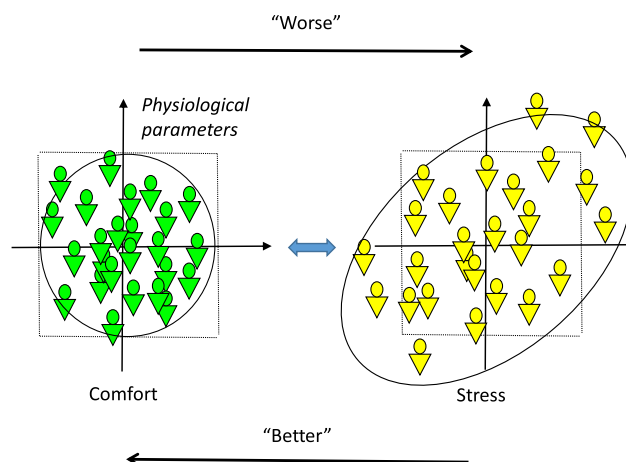


Fig. 2. Geometric representation of the effect. Each organism is represented as a data point in an n -dimensional vector space. The well-adapted organisms are not highly correlated and after normalization of scales to unit variance the corresponding cloud of points looks roughly as a sphere. The organisms under stress are highly correlated, hence in the same coordinates the data cloud looks like an ellipsoid with remarkable eccentricity. The largest diameter of this ellipsoid is larger than for the well-adapted organisms and the variance increases together with the correlations. The dotted square corresponds to the variance of attributes in the comfort zone normalized to unity.

1.2. Correlations and variability in the course of adaptation: the main directions of experimental research

First time, this effect was demonstrated on the lipid metabolism of healthy newborns [68]. We studied two groups: (i) babies born in the temperate belt of Siberia (the comfort zone) and (ii) in the migrant families of the same ethnic origin in a Far North city. (The families lived there in the standard city conditions.) A blood test was taken in the morning, on an empty stomach, at the same time every day. All data were collected in the summer. The correlation graphs are presented in Fig. 1.

The observed effect has a clear geometric interpretation (Fig. 2). A typical stress pattern can be expressed symbolically: Correlations $\uparrow$ (increase) and Variance $\uparrow$ (increase). This means that the answer to the following question is ambiguous: Do the organisms become more or less similar under stress? An increase in variance means more differences in attribute values, while an increase in correlations shows more similarity in the relationships between attributes.

After systematic testing on many groups and various types of stress, this effect was proposed as a method of screening of the population in hard living conditions (Far North city, polar expedition, army recruits, or groups that

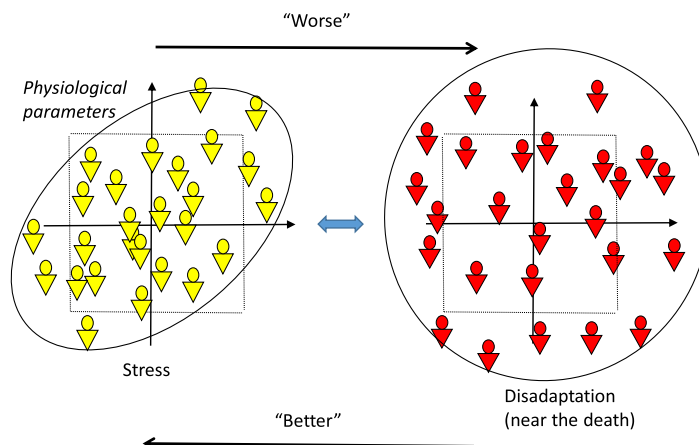


Fig. 3. Non-monotonicity of correlations as a function of stress intensity: when organisms are approaching death, correlations decay, while variance (or relative variation) continues to increase.

have changed their living conditions, for example) [146]. Network characteristics appear to be more informative indicators of stress than just attribute values. This was proven by many experiments and observations of groups of humans, mice, trees and grassy plants [70]. A simple but useful trick is to collect the ‘population’ (Fig. 2) from the same organism at different time moments. The effect was successfully applied in psychiatry and psychotherapy research [35,38,39]. In particular, correlations in the network of specific symptoms were used as a marker of major depression: Strong connections correspond to severe depression and deterioration of the patient’s condition, while weaker connections characterize stable state or recovery [35].

The correlation dynamics and Markov models of dyads ‘patients–psychotherapists’ was studied in [37–39,88], where a statistical mechanics – inspired quantitative approach to evaluation of effectiveness of psychotherapy was developed using analysis of correlation graph, PCA and cluster analysis. In particular, it was demonstrated that the outcome of psychotherapy can be qualitatively predicted (good or poor) on the basis of the correlation pattern between the coarse-grained verbal behavior of patients and psychotherapists [38]. A physiological approach to measuring of empathy between patients and psychotherapists was developed and tested based on correlations between patient and therapists heart rate and galvanic skin response [88]. The data-driven ‘synchrony index’ developed in these works meets the classical idea of synchrony in psychotherapy (see, for example, the review [91]).

Correlation indicators in attribute networks are useful in analyzing large socio-political systems. For example, the dynamics of 19 main public fears that have spread in Ukrainian society during a long stressful period preceding the Ukrainian economic and political crisis of 2014 were analyzed [139]. In the pre-crisis years, there was a simultaneous increase in the total correlation between fears (by about 64%) and their variance (by 29%).

In 1995, this effect (Fig. 2) was also described for the equity markets of seven major countries over the period 1960–1990 [99] and in 1997 for the twelve largest European equity markets after the 1987 international equity market crash [110]. These observations initiated development of econophysics [107]. Analysis of 2008 Global financial crisis was performed in [70]. There are attempts to use this effect in seismology to predict earthquakes [182]. It has also been selected as a part of the versatile critical transition anticipating toolkit [142].

In Section 2, “Correlation, risk and crisis”, we review the relevant data. The collection of examples presented in our previous works [70] has been significantly expanded by the findings of the last decade.

Most of the data, which we collected by ourselves or found in publications, support the hypothesis presented in Fig. 2. Therefore, special attention was needed to search for rear data that contradict this picture. The effect of non-monotonicity was revealed: when the adaptive abilities are exhausted and the organisms are close to death, the correlations decrease again, and the variance continues to grow (Fig. 3). Thus, near the fatal outcome, the dynamics of data can be represented as follows: Correlations↓ (decrease) and Variance↑ (increase). A simple but significant comment about dimensionless variables is needed. Correlations are dimensionless, but variances are not. In most physiological data we work with the variables that are positive by their nature (concentrations, activities, etc.). Near

the fatal outcome, the means may change drastically, therefore rescaling may be necessary. In such cases we should consider dimensionless *relative variation*.

When describing the effect, the idea of stress and the vague terms “better” and “worse” were used. What do these “better” and “worse” mean? Of course, when classifying specific clinical situations, we relied on the opinion of medical experts. All the experiments are unbiased in the following sense: the definitions of the “better–worse” scale were done before the correlation analysis and did not depend on the results of that analysis. Hence, one can state, that the expert evaluation of the stress and crisis can be (typically) reproduced by the formal analysis of correlations and variance.

However, a deeper analysis of the instantaneous evaluation of the ‘fitness of the individual’ is needed. The theoretical notion ‘fitness’ can be defined in the context of Darwinian selection but requires long time history for evaluation. Even the possibility to consider instantaneous individual fitness needs clarification. We consider this problem in Section 3.2.3, “The challenge of defining wellbeing and instantaneous fitness”.

1.3. Theoretical approaches ‘from a bird’s eye view

Such a general effect (Fig. 2) needs also a general theory for its explanation and for prediction of possible exclusions. The theoretical backgrounds were found in the factors-resource model. Adaptation is modeled as a process of redistributing resources to neutralize harmful factors (the lack of something necessary is also considered a harmful factor).

The general ‘factor-resource’ models were qualitatively discussed by Selye in his classical works about General Adaptation Syndrome (GAS) [147,148]. In analysis of his experiments, he used a physical (thermodynamic) analogy and proposed a universal adaptation resource – *adaptation energy* (AE). The literal interpretation of this analogy led to criticism of Selye’s approach, as the physical energy of adaptation was not revealed.

We considered AE as an internal coordinate on the ‘dominant path’ in the model of adaptation [72]. This approach demystifies the existence of a single resource variable for describing adaptive processes. The dominant path can appear as a manifold of slow motion [66] in dynamics of adaptation. Nevertheless, the additive behavior of this hidden coordinate of adaptation noticed by Selye [147,148] in discussion of his experiments may need additional explanation.

During last decade, the energetic concept of adaptation cost has been revived [94]. A general energy-speed-accuracy relationship was found between the rate of energy dissipation, the rate of adaptation and the maximum accuracy of the adaptation. The theory has been tested for the adaptation of chemoreceptors in the chemosensory system of *E. coli*. Analysis of adaptive heat engines that functions under scarce or unknown resources related the resources of adaptation to the prior information about the environment [2]. The hidden reservoirs of accumulated energy spent on adaptation are recognized as an essential component of adaptation to a non-stationary environment with uncertainty [2]. This recent resurrection of the thermodynamic foundations of Selye’s adaptation energy supports the idea of a single adaptive resource and factor-resource models. Simple but universal thermodynamic estimates of the cost of homeostasis are presented in Appendix A.

Section 3, “Selye’s thermodynamics of adaptation”, presents general models of the adaptation. In the general factor-resource models, the result of adaptation can be predicted, but not the dynamics of the adaptation process leading to this result.

The mechanism of adaptive redistribution of resource is determined by the principle of optimality, that is, by maximizing Darwinian fitness. The optimal mechanism of redistribution strongly depends on the interaction of harmful factors when they affect the organisms. Two opposite main cases are: the Liebig systems and the synergistic systems of factors. The Liebig systems obey the “Law of the Minimum” that states that growth and well-being are controlled by the limiting factor (the scarcest resource or the most harmful factor). In the synergistic systems harmful factors superlinear amplify each other. The optimal redistribution of the adaptation resource leads to a paradoxical result [69]:

- *Law of the Minimum paradox*: If for a randomly selected pair, (‘State of environment – State of organism’), the law of the minimum is valid (everything is limited by the factor with the worst value) then, after adaptation, many factors (the maximum possible number of them) are equally important. In the process of adaptation, such systems evolve towards breaking the law of the minimum.

- *Law of the Minimum inverse paradox*: If for a randomly selected pair ('State of environment – State of organism') many factors are equally important and superlinearly amplify each other then, after adaptation, a smaller amount of factors is important (everything is limited by the factors with the worst uncompensated values). In the process of adaptation, such systems evolve towards the law of the minimum).

The classification of interaction in pairs of resources was developed by Tilman [162,163]. He studied equilibrium in resource competition. Our theory of adaptation to different systems of factors uses very general properties of fitness functions and interactions between factors in combination with the classical idea of evolutionary optimality [68–70, 72]. Since fundamental works of Haldane (see [76]), the principles of evolutionary optimality, their backgrounds and applications are widely discussed in the literature [44,60,81,86,125,154]. Despite some limitations and criticism of the 'adaptationism' approach, it was decided that the fundamental contribution of optimality models is that they describe what organisms should do in particular instance [123]. Both formulated paradoxes are the consequences of this approach. In short, Law of the Minimum paradox means that the organism responds to situations with limitation by allocating the available regulatory resources to compensate for this limitation. This reaction reduces the influence of the limiting factor and may create a zone of co-limitation, where several factors are important.

For example, daphnids may respond to stoichiometrically imbalanced diets after ingestion [40], e.g. by down-regulating the assimilation of nutrients in excess and up-regulating the assimilation of limiting nutrients [78]. This response reduces limiting and leads to co-limitation [157].

The 'inverse paradox' also describes optimal regulation: if two harmful factors superlinearly amplify each other and are both important, then the optimal strategy is to allocate all the available regulatory resources on one of them and reduce its influence to the minimum. After that, the remaining factor will become limiting. Such interaction was discovered in the theory of resource competition for pairs of resources [162,163] and developed later to the general theory [70]. This effect was supported by some consequences like decrease of correlation in stress when the harmful factors are expected to amplify each other. Nevertheless, there is a need for direct experimental testing of this mechanism as it was done for appearance of colimitation in well-adapted systems.

An alternative approach to the dynamics of correlation graph under stress, based on the idea of *hidden bifurcations*, was proposed by several authors (see, for example, [32]). It assumes that there exists a dynamical system of physiological regulation. It may not be fully observable, but it does affect relevant variables. The observed increase of correlation and variance is considered as the result of the bifurcation in this dynamical system. Hypothetically, the disease is interpreted as a new attractor resulting from bifurcation (Sec. 3.6.1).

The *advection–diffusion model* for distribution of population in the space of physiological attributes assumes that the organisms perform random walk in the space of physiological attributes due to individual differences and various fluctuation of environment [135,136]. The area U of this walk is bounded. A harmful factor is represented in this model as a 'wind' that moves the population to the border of U . Dimension of the cloud of data reduces under this load because the datapoints under strong pressure are located near the boundary of U (Sec. 3.6.2). A question remains: how can this minor decrease in the dimension of a data cloud in multidimensional spaces explain a significant increase in correlations?

The adaptation models introduced and analyzed in this work exploit common phenomenological properties of the adaptation process: homeostasis (adaptive regulation), adaptation cost (adaptation resource), and the idea of optimization. The developed models do not depend on the particular details of the adaptation mechanisms. Models that do not depend on many details are very popular in physics, chemistry, ecology, and many other disciplines. They aim to capture the main phenomena. Thermodynamics is the Queen of these approaches, and it is no accident that language and ideas from thermodynamics are used at important stages of developing basic adaptation models.

Perhaps the most important challenge is assembling a holistic view of the problem of adaptation from the many different successful studies created over the past decades. This synthesis seems to be a difficult task, because the existing models even have different ontologies. Our vision is that simple and basic adaptation models can provide a useful workbench for this build.

In Conclusion and outlook we discuss the results and present the open problems and perspectives.

In Appendices, we provide formal details of the models that complement the overall picture.

2. Correlation, risk and crisis

In this Section, we review the main empirical findings about dynamics of correlation graph in the course of adaptation. Our goal is to demonstrate the universality of the effect for different ensembles of adaptable systems with particular attention to two additional questions:

1. Is it possible to convert the observed correlation graphs into early warning indicators of stress and crisis?
2. Are there controversies in the interpretations of reported observation?

2.1. Indicators of correlation strength

For a set of m attributes, $\{x^l | l = 1, \dots, m\}$, we have $m(m-1)/2$ correlation coefficients r_{kl} between them. The standard choice for numeric variables is the Pearson correlations coefficient:

$$r = \frac{\langle xy \rangle - \langle x \rangle \langle y \rangle}{\sqrt{\langle (x - \langle x \rangle)^2 \rangle} \sqrt{\langle (y - \langle y \rangle)^2 \rangle}} \quad (1)$$

where x_i, y_i are the sample values of the attributes x, y , $\langle \dots \rangle$ stands for the sample average value, $\langle x \rangle = \frac{1}{n} \sum_i x_i$, and n is the number of samples.

A correlation graph [16,68,177] is a graph whose vertices correspond to attributes, and edges connect vertices x^k, x^l with a correlation coefficient above the threshold $|r_{kl}| \geq \alpha$. It is standard practice to use multiple edge types (e.g. very strong correlations and moderately strong correlations - see Fig. 1.). Such graphs were intensively used in medical research [20,70,72,146]. For qualitative attributes, the mutual information can be used [21]. The correlation graph approach was developed in data mining [52,83,126,171] with applications in econophysics [28,70,96,121,122,155], genomics [117], social network analysis [175] and other areas.

To measure the intensity of correlations in general, the weight of the correlation graph is used:

$$G_{p,\alpha} = \left(\sum_{j>k, |r_{jk}|>\alpha} |r_{jk}|^p \right)^{\frac{1}{p}}. \quad (2)$$

This a p -weight of the α -correlation graph. The usual choice is $p = 1$ (i.e., the l_1 norm is used). The choice of α is more flexible and depends on the range of correlation coefficients. By default, either $\alpha = 0.5$ or $\alpha = 0.7$. Normalization or $G_{p,\alpha}$ ‘per pair of attributes’ can be useful for comparing different set of attributes.

For example, the average absolute value of the correlation coefficient per pair of attributes was systematically used to identify critical transitions in lung injury, liver cancer, and lymphoma cancer by analyzing microarray data for these three diseases [32]. In addition, for each selected feature group, the average absolute value of the internal correlation coefficient and the average absolute value of external correlations were used. The *composite index* was proposed

$$I = \frac{SD_d \cdot PCC_d}{PCC_o}, \quad (3)$$

where PCC_d is the average absolute value of the Pearson Correlation Coefficient (PCC) of the dominant group, PCC_o is the average absolute value of PCC between the dominant group and others, and SD_d is the average Standard Deviation (SD) of the dominant group.

The additional idea of [32] is in existence of the dominant group such that near crisis correlations and variance inside the group increase while correlations between this group and other attributes decrease. The composite index combines these three indicators ($PCC_d \uparrow$, $SD_d \uparrow$, and $PCC_o \downarrow$) in one: $I \uparrow$.

There are many correlation indicators based on the eigenvalues of the correlation matrix. If there are no correlations, then all eigenvalues are equal to 1. For strong correlations, the range of the eigenvalues is much wider. Two ideas are used: (i) estimating the number of major principal components [84] (i.e., the linear dimension of the data) and (ii) estimating the variability of the distribution of eigenvalues. Various definitions of principal components and non-linear generalizations are presented in [67,73]. Several methods for evaluation of the number of principal components to retain were compared in [22]. Recently, after testing of many definitions of data dimensionality, we can suggest that the most stable and convenient evaluation of the number of principal components for retaining is based on the

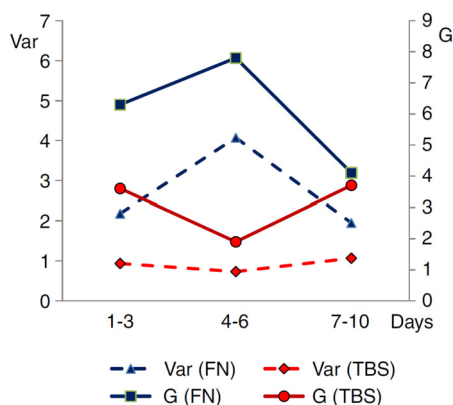


Fig. 4. The weight of the correlation graphs of eight lipid fractions (solid lines) and the variance (the relative variation) of these fractions (dashed lines) for healthy newborns born in the temperate zone of Siberia (TBS), and in families of migrants of the same ethnic origin in the city of the Far North (FN) during the first ten days after birth.

condition number of the reduced covariance matrix [65,113]: Intrinsic (linear) dimensionality of dataset is defined as the number of eigenvalues of the covariance matrix exceeding a fixed percent of its largest eigenvalue [53]. Some other indicators are presented, used and compared in [70].

The variability of the distribution of eigenvalues can be estimated by various methods, from standard deviations analysis to comparison with various ‘natural’ distributions such as the broken stick distribution [22,70] (the length of the pieces of a stick broken at $n - 1$ random points distributed uniformly and independently) or the distributions of eigenvalues of random matrices [149]. This latter approach has become popular in econophysics [5,130,133,167] because in typical situations where the sample size n and the number of attributes m are both large but comparable spurious correlations appear. Therefore, the empirical correlations should be compared not to the uncorrelated limit but to the fictitious correlations, which appear in $m \times n$ data matrices with independent, centralized, normalized and Gaussian matrix elements (the Wishart distribution [178]). If both $m, n \rightarrow \infty$ for constant ratio n/m then the limit distribution of eigenvalues has a simple density, the Marchenko–Pastur distribution [70,104,118] (for generalizations see [124]).

2.2. Humans, physiological data

2.2.1. Lipid metabolism of healthy newborn babies

Correlations of eight lipid fractions were studied in healthy newborns born in the temperate zone of Siberia (TBS), and in families of migrants of the same ethnic origin in the city of the Far North (FN) during the first ten days after birth [68]. The correlation graphs are presented in Fig. 1. On Fig. 4 we can see that the variance (Var) monotonically increases with the weight of the correlation graph. The number of babies: 1st-3rd days, TBS – 123 and FN – 100 babies; 4th-6th days, TBS – 98 and FN – 99 babies; 7th-10th days, TBS – 35 and FN – 29 babies.

2.2.2. Adaptation to change of climate zone

We studied people who moved to the Far North, six months after moving [18,19]. Two groups were compared: the test group of 54 people that had any illness during the period of short-term adaptation, and the control group, 98 people without illness during the adaptation period. We analyzed the activity of enzymes (alkaline phosphatase, acid phosphatase, succinate dehydrogenase, glyceraldehyde3-phosphate dehydrogenase, glycerol-3-phosphate dehydrogenase, and glucose-6-phosphate dehydrogenase) in leucocytes. The test group demonstrated much higher correlations between activity of enzymes than the control group (evaluated after 6 months at Far North), $G = 5.81$ in the test group versus $G = 1.36$ in the control group. For these groups, the dimensionless variance was compared: the enzyme activities were scaled to unit sample mean values, which is necessary, since the normal enzyme activity differs by orders of magnitude. For the test group the sum of these relative variations of the normalized enzyme activities was 1.204, and for the control group it was 0.388.

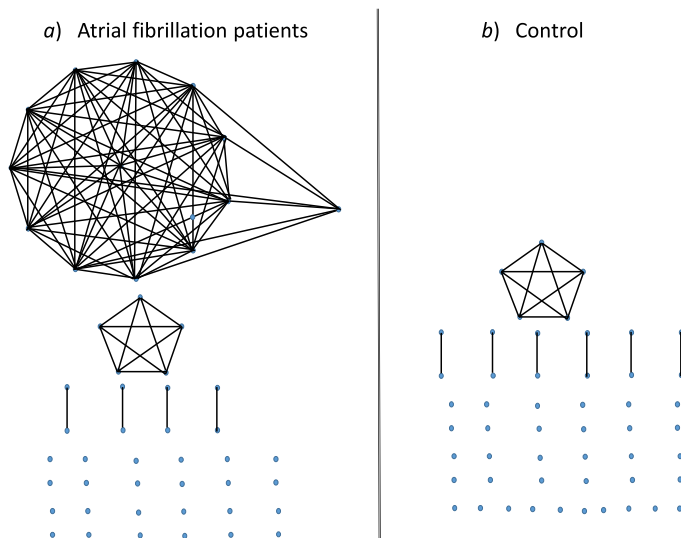


Fig. 5. The correlation graph for genes differentially expressed between patients with atrial fibrillation (a) and healthy control (b) (according to data analysis from [26,75]). The nodes correspond to different genes. Edges correspond to pairwise intra-class between-genes correlations with an absolute value greater than 0.90.

2.2.3. Gene regulation networks in atrial fibrillation

Increased of connectivity as a response to stressful condition was clearly demonstrated on the correlation network of the gene expression. The differences in gene expression profiles between atrial fibrillation patients and healthy controls are related to a very small part of the total data variability. Nevertheless, the factor analysis succeeds in discriminating patients from controls and extracting further genes involved in the pathology [25]. This approach identified groups of genes involved in structure and organization of cardiac muscle, and in inflammatory processes. Additionally, some genes that allow completing the transcriptomic deregulation picture of the pathology were detected. The correlation graph for the expression of the detected genes is presented for patients with atrial fibrillation in Fig. 5a and for healthy control in Fig. 5b (according to data analysis from [26,75]). Edges in (Fig. 5) correspond to high correlations with an absolute value of the Pearson correlation coefficient greater than 0.90.

2.2.4. Gene network analysis for muscular dystrophy

In some special cases, the dynamics of the correlation graph contradicts the typical picture shown in Fig. 2 (but still supports the idea that the correlation graph can be a better indicator of critical transitions than the attribute values themselves). A seminal case study was performed on the gene expression data of skeletal muscle from Duchenne muscular dystrophy (DMD) patients [27]. The genome-wide gene expression profiling of skeletal muscle from 22 DMD children were collected with 14 age-matched controls [128]. The DMD patients were at the initial or ‘presymptomatic’ phase of the disease. The control group consisted of patients who came to the hospital with a suspect metabolic disorder that was not confirmed by biochemical and histopathological studies. The data are available [127].

The differences in gene expression profiles between DMD patients and controls are related to a very small part of the data variability [27]. Principal Component Analysis (PCA) helps to overcome this difficulty and to find the relevant signals. The first two principal components show no difference between the two groups. The first principal component explains more than 98% of the total variance. Nevertheless, in the projection on the third principal component, the gene expression profiles of the patients with DMD are well separated from the control group. The gene having the highest score for the third PC was dystrophin (as expected). Correlation network of 100 genes with the highest scores in the third principal components were analyzed.

Correlation was considered significant when above the threshold value, obtained by the surrogate data analysis (0.84). The activity of important dystrophin gene was not correlated with other genes, both in DMD patients and in control. There were 85 significant connections in the DMD patients correlation graph and 133 connections in the control [27]. Thus, in general, the connectivity of the correlation graph is lower for the DMD patients.

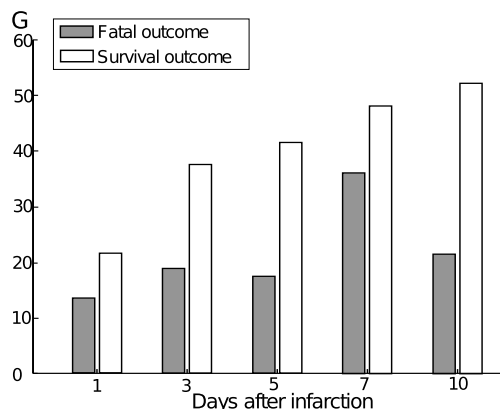


Fig. 6. Dynamics of weight of the correlation graphs of echocardiography-derived variables, parameters of central hemodynamics, biochemical parameters, and also leucocytes during 10 days after myocardial infarction for two groups of patients: for the survival outcome and for the fatal outcome. Here G is the sum of the strongest Pearson's correlation coefficients $|r_{ij}| > 0.4$, $i \neq j$ [158].

In more detail, there were two subnetworks in the control group, A and B [27]. The subnetwork A consisted of genes that encode various hemoglobins. The group of DMD patients had practically the same subnetwork. The subnetwork B for the control group consisted of molecules involved in the extracellular matrix and cytoskeleton organization. For DMD patients, this subnetwork was much less correlated and split into three disconnected subnets.

Nevertheless, for the DMD patients there appeared a special new connected component C of correlated genes which were not correlated in the control group. This subnetwork consisted of genes involved in muscle development. If we focus on this small muscle development network, then the dynamics of correlations still follows Fig. 2.

Thus, this case study demonstrated that dynamics of correlations for illness and critical transitions can be multidirectional: some subnetworks may become less correlated, while correlations in some other subnetworks may increase. For such cases, “the devil is in the details” of choosing subnetworks for correlation analysis.

For example, the same procedure applied to the mitochondrial network genes in the skeletal muscle of amyotrophic lateral sclerosis patients demonstrates the classical behavior (Fig. 2). There are higher levels of correlation among genes, whose function are aberrantly activated during the progression of muscle atrophy [11]. The genes in this network seem to reflect the perturbation of muscle homeostasis and metabolic balance occurring in affected individuals.

2.2.5. Myocardial infarction and ‘no-return’ points

The dynamics of correlations between physiological parameters after myocardial infarction was studied in [158]. For each patient (more than 100 people), three groups of parameters were measured: echocardiography-derived variables (end-systolic and end-diastolic indexes, stroke index, and ejection fraction), parameters of central hemodynamics (systolic and diastolic arterial pressure, stroke volume, heart rate, the minute circulation volume, and specific peripheral resistance), biochemical parameters (lactate dehydrogenase, the heart isoenzyme of lactate dehydrogenase LDH1, aspartate transaminase, and alanine transaminase), and also leucocytes. Two groups were analyzed after 10 days of monitoring: the patients with a lethal outcome, and the patients with a survival outcome (with compatible amounts of group members). These groups do not differ significantly in the average values of parameters and are not separable in the space of measured attributes. Nevertheless, the dynamics of the correlations in the groups are essentially different. For the fatal outcome correlations were stably low (with a short pulse at the 7th day), for the survival outcome, the correlations were higher and monotonically grew. This growth can be interpreted as return to the “normal crisis” (the left position in Fig. 3). Topologically, the correlation graph for the survival outcome included two persistent triangles with strong correlations: the central hemodynamics triangle, minute circulation volume – stroke volume – specific peripheral resistance, and the heart hemodynamics triangle, specific peripheral resistance – stroke index – end-diastolic indexes. The group with a fatal outcome had no such persistent triangles in the correlation graph.

In the analysis of fatal outcomes for oncological patients [103] and in special experiments with acute hemolytic anemia caused by phenylhydrazine in mice [132] (see also Sec. 2.4.1) one more effect was observed: the correlations increased for a short time before death, and then fell down (see also the pulse in Fig. 6). This pulse of the correlations

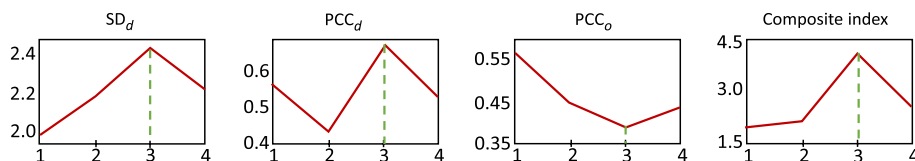


Fig. 7. Early warning signals for the hepatic lesion by chronic hepatitis B. SD_d is the average standard deviation of gene expression in DNB, PCC_d is the average Pearson correlation coefficient per pair of gene expressions in DNB, PCC_o is the average Pearson correlation coefficient between gene expression in DNB and their external neighborhood gene expression. The composite indicator is given by (3). The vertical dashed lines indicate the pre-disease state. The figure is drawn using data from [32].

(in our observations, usually for one day, which precedes the fatal outcome) is opposite to the major trend of the systems in their approach to death. We cannot claim the universality of this effect and it requires special attention.

2.2.6. Genomic data about hepatic lesion by chronic hepatitis B

Significant numbers of genes were up- or down-regulated in the liver with chronic hepatitis compared with normal liver. The dynamics of correlation and variance in gene expression at the disease and pre-disease stages were analyzed in [32] using earlier data on differential gene expression [80]. There were 12 patients with disease and 6 patients in control. A subnetwork of the dynamical network biomarker (DNB) was identified. According to [32], the variance of DNB gene expressions and correlations between them achieve their maxima at the tipping point of the pre-disease before it is transformed into the disease. Correlations between DNB gene expressions and the expressions of the genes that do not belong to DNB is minimal at this point.

The dynamics presented in Fig. 7 partially supports the hypothesis formulated above (Fig. 2): in the pre-disease stress, both correlation and variance increase. The question appeared about the behavior of correlation and variance after the tipping point #3 (Fig. 7). According to the observation illustrated in Fig. 3, the non-monotonicity of correlations as a function of stress intensity is an expected effect but for a really high stress load (see also [70]).

Another difference is in the dynamics of the SD. According to [32], SD_d decreases after the tipping point (Fig. 7). According to our observations (in different case studies), on the contrary, the average variance continues to increase after the correlations begin to decrease. Differences (apart from different theoretical models) can be caused by different data preprocessing. We used a dimensionless relative SD (normalized to the unit mean value of the positive attributes). This normalization makes sense because (i) the values of various attributes can differ by orders of magnitude, and without normalization, many SD inputs may practically disappear, and (ii) these values are positive numbers in most cases. The work [32] used the CD “as it is”. The difference in the emerging dynamics may be caused by a possible decrease in gene expression in DNB after the tipping point. This issue requires more attention in the future.

2.2.7. Therapy of obesity

The study was conducted on patients with different levels of obesity [150,170]. The study reported in [150] included 235 patients aged 34 to 79 years, suffering from 1-3 degrees of obesity. All patients, depending on the degree of obesity and the nature of concomitant pathology, were divided into 3 groups. The first group of the study included patients with the first degree of obesity. The second group consisted of patients suffering from the second and third degrees of obesity in combination with functional disorders of various organs and systems of the body (dyskinetic disorders of the digestive system, arterial hypertension of the first degree, asthenic syndrome). The third group included patients who had organic lesions (peptic ulcer, arterial hypertension of the third degree, patients after a heart attack, stroke, etc.) against the background of the second and third degrees of obesity.

All patients received a traditional course of treatment for 60 days, aimed at reducing body weight and correcting metabolic and organ disorders. Treatment of patients of the first group was limited to diet therapy, the second group - with the additional statin medication (Crestor), the third group — with the additional drug therapy against comorbidities. The study included the following indicators: body mass index, fat mass, lean mass, total water, total cholesterol, high-density cholesterol, low-density cholesterol, creatinine, and triglycerides. The weight of the correlation graph G (the Table 1) was originally high and monotonically dependent on the level of sickness. It decreased during therapy.

Table 1

The weight G of correlation graph between attributes for patients with different level of obesity before and after treatment.

	Before treatment	After treatment
Group 1	8.24	7.33
Group 2	10.74	8.09
Group 3	13.41	11.72

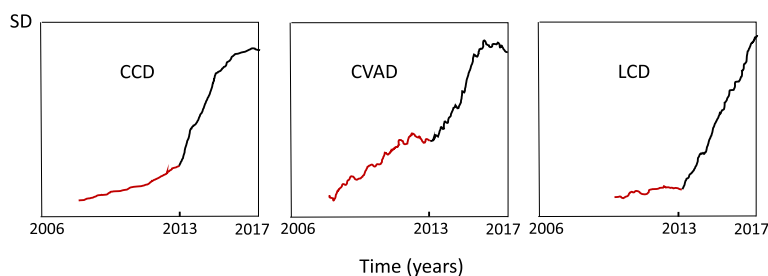


Fig. 8. Dynamics of CD of daily visits to a hospital in Nanjing City, China for CCD, CVAD and LCD during 2006–2016. The figure qualitatively reproduces CD dynamics from [173]. All the observed tipping points are close to the beginning of 2013.

2.2.8. Critical transitions in cardiopulmonary population health related to air pollution

Fluctuations in time series of hospital visits were analyzed to identify a tipping point in the health status of the population [173]. Fluctuations were quantified by standard deviations (SD) and autocorrelations at lag-1 (AR-1) and were used as potential early warning indicators. The hospital visits from a hospital in Nanjing City, China during ten years (2006–2016) were studied for the following cardiopulmonary diseases: cerebrovascular accident disease (CVAD), coronary artery disease (CAD), chronic obstructive pulmonary disease (COPD), lungcancer disease (LCD), and the grouped category of the respiratory system diseases (RES) with of the cardio-cerebrovascular system diseases (CCD). All these diseases are affected by air pollution.

The effect of critical transitions with a fast increase in SDs was clearly demonstrated for three diseases, CCD, CVAD and LCD (Fig. 8). From the data reported in [173], it seems likely that the increase in CDs for these diseases is significantly faster than the increase in visits (smoothed by a moving average filter).

The observed critical transition was interpreted as a health impact of air pollution. Of course, the deeper analysis should combine the data about hospital visits with pathophysiological clinical studies. This is a standard problem of labeling critical transitions found using indirect indicators: what critical transition is observed (for example, in the population health, or in hospital services, or in the structure and behavior of the population)?

2.2.9. Network rewiring in cancer: applications to melanoma

A detailed analysis of rewiring correlation networks of gene expressions was presented in [41] for melanoma. The authors argue that an approach to analyzing molecular pathophysiological processes based only on gene expression values cannot succeed in predicting the potential effect of a drug because it does not solve the important problem of the relationship between genes. It is shown that differences in the average expression may be insignificant, while the difference between correlations gives statistically significant p -values.

The differences in the topology of rewired networks can be enormous because cancer cell lines may have undergone numerous changes in genetic network correlation and expression patterns. Gene centrality profiles were used to quantify these differences for future use in drug target determination. These profiles characterize the connections and their strength from a gene to other genes. It is assumed that genes that play a more central role in subsets of genes within a broader relevant network are better targets for drugs in a disease state. The authors claim that their results shed light on the understanding of the molecular pathology of melanoma, as well as on the choice of treatment.

2.3. Humans, psychological data

Correlation graph analysis has proven to be useful in psychiatric and psychological research. Here, instead of a population or a group of different organisms shown in Fig. 2, a ‘population’ of states of one person at different points in time is considered. For this section we selected one recent achievement in this area, where individual patients are characterized by their own personal network with unique architecture and resulting dynamics [35]. For correlation analysis of dyads ‘patients–psychotherapists’ we refer to [37–39,88].

2.3.1. Dynamics of depression

The network of symptoms of major depression was analyzed in [35]. The vertices of this network correspond to 14 symptoms: insomnia, fatigue, concentration problems, depressed mood, feelings of self-reproach, etc. The main hypothesis is that symptoms can induce each other and individuals differ in the strength of the connections between certain symptoms linked in the network. The authors of [35] went beyond a simple correlation graph analysis and proposed a network model of symptom dynamics that resembles a probabilistic recurrent neural network. This model appears to be very general and can be applied to dynamical networks of attributes of many diseases.

Each (i th) symptom is characterized by a Boolean variable X_i with an obvious interpretation: $X_i = 1$ means that the i th symptom is ‘on’ (active) and $X_i = 0$ means that it is ‘off’ (inactive). We observe the patient in discrete time with a constant time step, $t = 1, 2, 3 \dots$. The probability that the i th symptom is active in the time moment $t + 1$ depends on the symptoms and external stress parameters at time t . The simple logistic regression was used:

$$\mathbf{P}(X_i(t+1) = 1) = \frac{1}{1 + \exp(b_i - A_i(t))},$$

where $A_i(t)$ is the signal from the i th ‘input summator’:

$$A_i(t) = \sum_j W_{ij} X_j(t) + S_i(t),$$

b_i is the threshold, W_{ij} is the weight matrix, and $S_i(t)$ is the stress parameter that may be specific for the i th symptom.

There are many methods for evaluation of the weight matrix and thresholds from data. An *IsingFit* function from open source R package was used in [35] under assumption that there was no ‘autocatalysis’ (no self-excitation of symptoms in the network). This assumption, of course, is not necessary. The weight matrix W_{ij} does not have to be symmetric because it describes effect of $X_j(t)$ on the i th symptom at the next time step, $X_i(t+1)$. For strongly connected networks (large values of W_{ij}), the long self-sustained major depression is possible even without exposure to external stress. On the contrary, in weakly connected networks, the depression decays without external stress if the thresholds are not too small.

This model sheds light on the mechanisms of development of major depression, explaining and predicting the dynamics of mutual induction of symptoms, for example, insomnia increases fatigue, which leads to problems with concentration, etc. The critical transitions to major depression are described using the connectivity of the network of symptoms. But the main result of [35] is, from our point of view, the introduction of a new class of simple and identifiable disease models with dynamic symptoms that can appear and disappear. We should also mention that the weight matrix W_{ij} is, in its essence, the correlation matrix, and the proposed model demonstrates how the correlation graph can be used for dynamic modeling of interactions between varying attributes.

2.4. Mice

Most data demonstrate that under stress and at the onset of illness the correlation and the variance increase (with an important remark that the choice of vertices for constructing a relevant graph can be a non-trivial task). This is just a beginning of the ‘medical history’. The question of a system near the fatal outcome is also important: what does the dynamics of Y correlations look like near the point of no return?

This problem about the dynamics of correlations near the point of no return was studied by analysis of fatal outcomes in oncological [103] and cardiological [158] clinics, and also in special experiments with acute hemolytic anemia caused by phenylhydrazine in mice [132] and phosgene inhalation lung injury [32,145]. The main result here is: when approaching the no-return point, correlations destroy (G decreases).

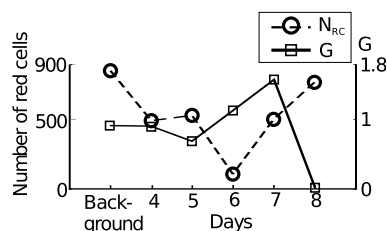


Fig. 9. Adaptation and disadaptation dynamics for mice after phenylhydrazine injection. Weight G of the correlation graph and number of erythrocytes N_{RC} .

There exists no formal conventional criterion to recognize the situation “on the other side of crisis”. Nevertheless, the labeling of such situation is needed. The common sense “general practitioner point of view” [57,58] can help in the labeling. From this point of view, the situations described below are close to fatal outcome and should be considered as ‘the other side of crisis’: the acute hemolytic anemia caused by phenylhydrazine in mice with lethal outcome and phosgene inhalation lung injury with high mortality rate.

2.4.1. Destroying of correlations “on the other side of crisis”: acute hemolytic anemia in mice

This effect was demonstrated in special experiments [132]. Acute hemolytic anemia caused by phenylhydrazine was studied in CBAXlac mice. After phenylhydrazine injections (60 mg/kg, twice a day, with interval 12 hours) during first 5-6 days the number of red cells decreased (Fig. 9), but at the 7th and 8th days this number increased because of spleen activity. After 8 days most of the mice died. Dynamics of correlation between hematocrit, reticulocytes, erythrocytes, and leukocytes in blood is presented in Fig. 9. An increase in G preceded an active adaptation response, but G decreased to zero before death. The number of red cells increased also at the last day.

2.4.2. Phosgene inhalation lung injury

Carbonyl chloride (phosgene) poisoning leads to life-threatening pulmonary edema and irreversible acute lung damage. A genomic approach was used to investigate the molecular mechanism of phosgene-induced lung damage in mice [145]. CD-1 male mice were exposed whole body to either air (control) or phosgene for 20 min. Lung tissue was collected from air- or phosgene-exposed mice at nine time moments: 0, 0.5, 1, 4, 8, 12, 24, 48, and 72 h postexposure. The strongest physiological effects occur within the first 8 hours after exposure. They lead to increased levels of the BALF protein, severe pulmonary edema, and increased mortality [145]. In particular, after 12 hours, the mortality rate reached 50% – 60%, and after 24 hours it was 60% – 70%. Oligonucleotide microarrays were used to determine global changes in gene expression in the lungs of mice after exposure to phosgene. These data reveal biological processes involved in phosgene toxicity including GSH biosynthesis (previously known), angiogenesis, cell death, cholesterol biosynthesis, cell adhesion and regulation of the cell cycle. PCA was applied to determine the greatest sources of data variability. The genes most significantly changed as a result of phosgene exposure were identified and categorized based on molecular function and biological process [145].

These data were analyzed further in [32]. A relevant subnetwork of the dynamical network biomarker (DNB) was identified (220 genes and 1167 links). Dynamics of the standard deviation and correlations in the DNB, correlations between the DNB and other molecules, and the composite index were presented in Fig. 10.

2.5. Plants

2.5.1. Grassy plants under trampling load

The effect exists for plants too. The grassy plants in oak tree-plants are studied [87]. For analysis the fragments of forests are selected, where the densities of trees and bushes were the same. The difference between those fragments was in damaging of the soil surface by trampling. Tree groups of fragments are studied:

- Group 1 – no fully destroyed soil surface;
- Group 2 – 25% of soil surface are destroyed by trampling;
- Group 3 – 70% of soil surface are destroyed by trampling.

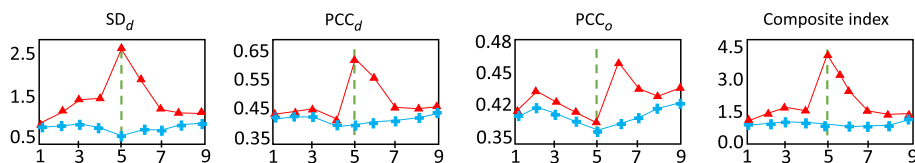


Fig. 10. Early-warning signals for phosgene-induced lung damage in mice. The horizontal axis represents the time period t . Red triangles correspond to the case group, while blue crosses represent the control group. The time scale is not uniform, and the 9 time moments are: 0, 0.5, 1, 4, 8, 12, 24, 48, and 72 h postexposure. Here, SD_d is the average standard deviation of gene expression in DNB, PCC_d is the average Pearson correlation coefficient per pair of gene expressions in DNB, PCC_o is the average Pearson correlation coefficient between gene expression in DNB and their external neighborhood gene expression. The composite indicator is given by (3). The vertical dashed lines indicate the pre-disease state. The figure is drawn using data from [32,145]. Clearly, the composite indicator starts increasing significantly from the 4th time period (4h), and reaches the peak in the 5th time period (8h), while the OPCC decreases. This fact was interpreted by [32] as a demonstration that the pre-disease state is located near the 5th time period (between the fourth and the fifth sampling points), and the phase transition into the disease state occurs after the 5th time moment.

Table 2

Weight G of the correlation graph for different grassy plants under various trampling load.

Grassy Plant	Group 1	Group 2	Group 3
Lamiastrum	1.4	5.2	6.2
Paris (quadrifolia)	4.1	7.6	14.8
Convallaria	5.4	7.9	10.1
Anemone	8.1	12.5	15.8
Pulmonaria	8.8	11.9	15.1
Asarum	10.3	15.4	19.5

The studied physiological attributes were: the height of sprouts, the length of roots, the diameter of roots, the amount of roots, the area of leaves, the area of roots. Results are presented in Table 2.

2.5.2. Scots pines near a coal power station

The impact of emissions from a heat power station on Scots pine was studied [151]. For diagnostic purposes the secondary metabolites of phenolic nature were used. They are much more stable than the primary products and hold the information about past impact of environment on the plant organism for longer time.

The test group consisted of Scots pines (*Pinus sylvestric* L) in a 40 year old stand of the II class in the emission tongue 10 km from the power station. The station had been operating on brown coal for 45 years. The control group of Scots pines was from a stand of the same age and forest type, growing outside the industrial emission area. The needles for analysis were one year old from the shoots in the middle part of the crown. The samples were taken in spring in bud swelling period. Individual composition of the alcohol extract of needles was studied by high efficiency liquid chromatography. 26 individual phenolic compounds were identified for all samples and used in analysis.

No reliable difference was found in the test group and control group average compositions. For example, the results for Proantocyanidin content (mg/g dry weight) were as follows:

- Total 37.4 ± 3.2 (test) versus 36.8 ± 2.0 (control);

Nevertheless, the variance of compositions of individual compounds in the test group was significantly higher, and the difference in correlations was huge: $G = 17.29$ for the test group versus $G = 3.79$ in the control group.

2.5.3. Drought stress response in sorghum

The drought-specific subnetwork for sorghum was extracted and described in detail in [179]. First, authors found 14 major drought stress related hub genes (DSRhub genes). These genes were identified by combination of various approaches and analysis of data from different sources like analysis of gene expression, regulatory pathways, sorghumCyc, sorghum protein-protein interaction, and gene ontology. Then, investigation of the DSRhub genes led to revealing distinct regulatory genes such as ZEP, NCED, AAO, MCSU and CYP707A1. Several other protein families

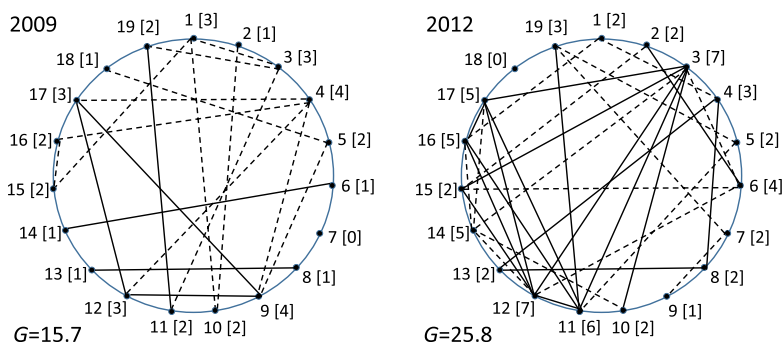


Fig. 11. The correlation graph of the prevalence of 19 main fears in regions of Ukraine for 2009 ($G = 15.7$) and 2012 ($G = 25.8$). The vertices with the PCC $|r_{ij}| > 0.7$ are connected. The solid line means that $r_{ij} > 0.7$ (strong positive correlation), the dashed line means that $r_{ij} < -0.7$ (strong negative correlation). The numbers in brackets are the orders of the vertices (the number of links). The data from [139] are used.

were found to be involved in the response to drought stress: aldehyde and alcohol dehydrogenases, mitogene activated protein kinases, and Ribulose-1,5-biphosphate carboxylase.

This analysis resulted in construction a drought-specific subnetwork, characterized by unique candidate genes that were associated with DSRhub genes [179]. Authors analyzed connections in this network and constructed pathway cross-talk network for 69 significantly enriched pathways.

In particular, they found that the stress increases the overall association and connectivity of the gene regulation network, even if different stress conditions, such as desiccation, salt and oxidative stresses in addition to cold or heat, occur in combination. For more detail and specific presentation of the drought-specific subnetwork of genes and pathway cross-talk network we refer to the original work [179].

2.6. Social systems

Anticipating critical changes in social systems is a big problem. Measuring ‘social stress’ and tracking its dynamics is a top priority in public administration. Unfortunately, very often social stress is measured post-factum after devastating social upheavals. The problem of reliably measuring social stress before critical events is still open. This problem was attacked recently using the analysis of correlation graph and variance under basic hypothesis presented in Fig. 2 [139]. In particular, they clearly demonstrate that the dramatic events in Ukraine in December 2013 and February 2014 were preceded by a monotonous increase in social tension during several years. This was not just a sudden ‘social explosion’.

For analysis, the data collected by the Sociological Pollster Group ‘Rating’ were used. In 2009–2012, this group conducted several surveys, in which respondents named three most important threats faced by Ukraine. Data were aggregated for four main regions of Ukraine: for six geographical regions of the country: (a) *The west* (Chernivtsi, Ivano-Frankivsk, Lviv, Rivne, Ternopil, Volyn and Zakarpattia provinces); (b) *The centre* (Cherkasy, Khmelnytsky, Kirovohrad, Poltava and Vinnytsia provinces); (c) *The north* (Chernihiv, Kyiv, Sumy and Zhytomyr provinces); (d) *The south* (Crimea, Kherson, Mykolaiv and Odessa provinces); (e) *The east* (Dnipropetrovsk, Kharkiv and Zaporizhia provinces); and (f) *Donbas* (Donetsk and Luhansk provinces).

Nineteen main fears were identified for analysis: (1) Economic regress, (2) Rise in unemployment, (3) Depreciation of national currency, (4) Arbitrary rule, (5) Degeneracy of population, (6) Health services’ worsening, (7) Environmental accidents, (8) Rise in crime, (9) Mass exodus, (10) Demographic crisis, (11) Schism of the state, (12) Losing sovereignty, (13) Civil war, (14) Education services’ worsening, (15) Losing control over the gastransport system, (16) Coup d’etat, (17) Military aggression from Russia, (18) Terrorism, (19) Military aggression from West.

For each region, a 19-dimensional vector of the prevalence rates of the public fears was identified. The correlations between fears were analyzed. For this small sample (6 regions) the probability to find the correlation coefficient between two given attributes $r > 0.7$ by chance is $p \approx 0.06$. The l_1 weight of the correlation graph G (2) was defined with the threshold $\alpha = 0.7$. The connectivity of the correlation graph definitely increased between 2009 and 2012 (Fig. 11). The variance of the prevalence of various fears also increased [139].

This simple analysis can serve as a prototype measurer of social stress. It seems promising to combine the correlation analysis of social fears with a dynamic model of network of symptoms developed for individual depression [35] (see Subsection 2.3.1).

2.7. Markets and finance

Intensive study of transformation of correlation graph under stress in physiology was started at the 1980s [68,146]. Some years later, in 1995, the seminal work of Longin and Solnik [99] demonstrated that similar effects exist in market dynamics: correlations in equity market are also non-constant and near crisis correlations increase.

During 25 years after this work, the effect of increase of correlations in crisis was demonstrated for many financial time series. It belongs now to the basis of econophysics [5,107,130,167], and analysis of correlation graphs is an important instrument for investigation of financial market.

Special attention was paid to the analysis of crises [70]. Here are some important and well-represented examples. Two phases of the Asian crisis were identified: (1) an increase in correlation (contagion) and (2) a continued high correlation (herding). A shift in variance during the crisis period was also detected. These effects have raised doubts about the benefits of international portfolio diversification [33]. Sudden and gradual changes in correlation between stocks, bonds and commodity futures returns were estimated. Most correlations start the 1990s at a low level but increased around the early 2000s and reached peaks during the crisis. Diversification benefits to investors were significantly reduced [153]. The crisis has put pressure on emerging markets in the Middle East and North Africa (MENA) region in particular. Correlations between MENA stock markets and the more developed financial markets, and the intra-regional financial linkages between MENA countries' financial markets were analyzed in [116]. A multiscale correlation analysis of the stock market during the global financial crisis was carried out for the G7 and BRIC countries and a new multiscale correlation contagion statistic test was developed for to support decision-making on the global portfolio diversification in crisis [174]. Partial correlation financial networks were introduced and applied to analysis of market crashes [112]. Structural breaks of eight national stock markets associated with the Asian and the Global financial crises were analyzed. Significant cross effects and long range volatility dependence were revealed [85]. Analysis of Russian banking and monetary system in global crisis was performed through the analysis of correlation graph and an estimation of conjugacy of monetary and banking policy. Hidden internal patterns were detected in this conjugacy and hidden systemic crisis of Russian banking and monetary system was revealed in 2007 data [131]. This crisis can be indicated by a sudden increase ('explosion') in the values of variance.

The correlation graph approach works also for analysis of a single company. The method of the structure and indicators analysis of a company business processes based on the analysis of correlations in historic series of expenses is developed [105]. This method was used to identify periods of stress in the company, to optimize and reengineer the management process and to allocate and reallocate resources between company functions. It can be also used by audit companies to identify cheating in a company's financial statements. Correlation graph analysis also helps to uncover fraudulent behavior hidden in the business distribution channels, where colluding partners enter into fake big deals in order to obtain lower product prices – behavior that is considered extremely detrimental to the sales ecosystem. [101].

2.7.1. Correlation anatomy of global financial crisis

In this Subsection, we give just one illustration and refer to special literature for more detail. Let us illustrate the general rule (Fig. 2) with the dynamics of the stock market, close to the famous financial crisis of 2008, following our work [70]. We used the daily closing values over the time period 03.01.2006 – 20.11.2008 for companies that are registered in the FTSE 100 index (Financial Times Stock Exchange Index). The FTSE 100 is a weighted index representing the performance of the 100 largest UK-domiciled blue chip companies. Thirty companies that had the highest value of the capital on Jan 1, 2007, were selected for the analysis of correlations in financial time series. The companies, the corresponding abbreviations, and the business type are listed in Table 3. Data for these companies were available on the Yahoo!Finance.

Data presented in Figs. 12, 13 demonstrate how correlations of the daily closing values in the sliding window (20 days) increase in crisis. Dynamics of correlations can be used as an indicator of critical transitions. Moreover, sometimes the changes in correlation occur earlier than in index and there is a chance to use the correlation graph as an early warning signal. The correlation graphs can be used for detailed analysis of the crisis spreading. Thus, we can see how the strong correlation appeared in the financial sector and then propagated to other types of business. The

Table 3
Thirty companies from the FTSE 100 index chosen for analysis.

Number	Business type	Company	Abbreviation
1	Mining	Anglo American plc	AAL
2		BHP Billiton	BHP
3	Energy (oil/gas)	BG Group	BG
4		BP	BP
5		Royal Dutch Shell	RDSB
6	Energy (distribution)	Centrica	CNA
7		National Grid	NG
8	Finance (bank)	Barclays plc	BARC
9		HBOS	HBOS
10		HSBC HLDG	HSBC
11		Lloyds	LLOY
12	Finance (insurance)	Admiral	ADM
13		Aviva	AV
14		LandSecurities	LAND
15		Prudential	PRU
16		Standard Chartered	STAN
17	Food production	Unilever	ULVR
18	Consumer goods/food/drinks	Diageo	DGE
19		SABMiller	SAB
20		TESCO	TSCO
21	Tobacco	British American Tobacco	BATS
22		Imperial Tobacco	IMT
23	Pharmaceuticals (inc. research)	AstraZeneca	AZN
24		GlaxoSmithKline	GSK
25	Telecommunications	BT Group	BTA
26		Vodafone	VOD
27	Travel/leisure	Compass Group	CPG
28	Media (broadcasting)	British Sky Broadcasting	BSY
29	Aerospace/defence	BAE System	BA
30		Rolls-Royce	RR

trajectory of the reverse process is very different as we can from the comparisons of cliques and strongly correlated clusters in Fig. 12b and c (especially in the correlation graphs for 17/05/2006 and 27/06/2006). The asymmetry between the ups and downs of the financial market was also noticed when analyzing the empirical financial correlation matrix of 30 companies included in Deutsche Aktienindex (DAX) [43].

2.8. What can we learn from examples?

- The correlation graph and variance are sensitive indicators of stress and illness.
- In the initial stages of stress and illness, correlations and variances tend to increase.
- This effect could be sensitive to selection of attributes for analysis and special procedures are needed for selection of relevant attributes. Correlations between relevant and irrelevant attributes may even decrease, while the dynamics of correlations between irrelevant attributes may be not as clearly defined.
- At the disadaptation stage and near the fatal outcome the correlation may decrease (the effect of ‘points of no return’). The dynamics of variance near the ‘points of no return’ is reported controversially: in some experiments the variance increases, while in others it decreases, which may be caused by different normalization and dynamics of the means. The hypothesis is that the relative variation should decrease.

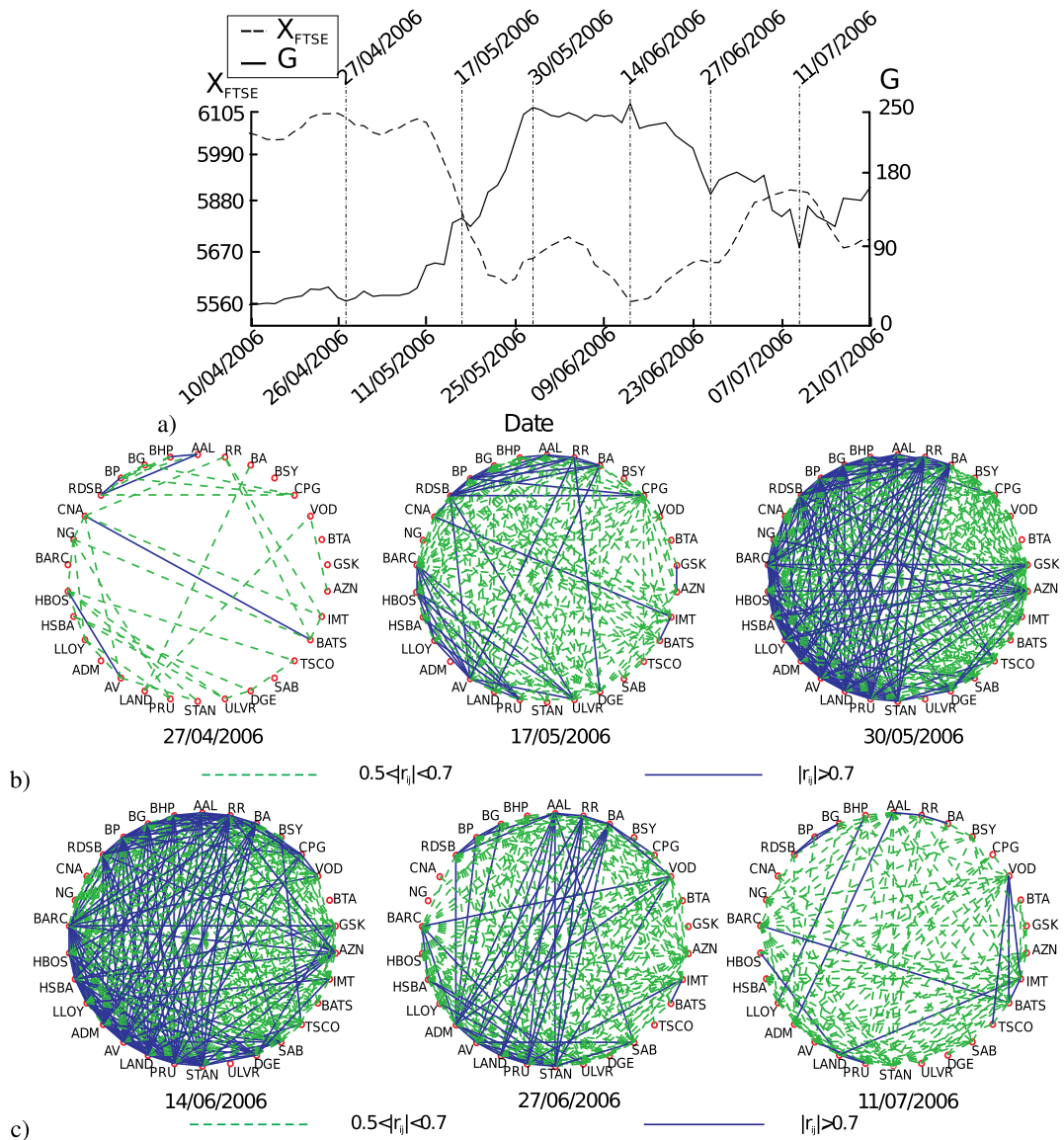


Fig. 12. Correlation graphs for six positions of sliding time window on interval 10/04/2006 - 21/07/2006. a) Dynamics of FTSE100 (X_{FTSE} , dashed line) and of $G = \sum_{j>i, |r_{ij}|>0.5} |r_{ij}|$ (solid line), vertical lines correspond to the points that were used for the correlation graphs. b) The correlation graphs for the first three points, FTSE100 decreases, the correlation graph becomes more connective. c) The correlation graphs for the last three points, FTSE100 increases, the correlation graph becomes less connective.

- The effect is observed in the ensembles of similar systems that adapt to the load of the same factors (organisms under load of harmful factors, enterprises in adverse conditions, etc.).
- These ensembles of similar systems can be collected from the history of one system at different time periods.
- Construction of early warning indicators of stress, illness and various crises on the basis of correlation graph and variance of relevant attributes is possible.
- For each class of systems, family of harmful factors and type of crisis, relevant data collection and a special data mining procedure are required to develop a stress and crisis metrology.

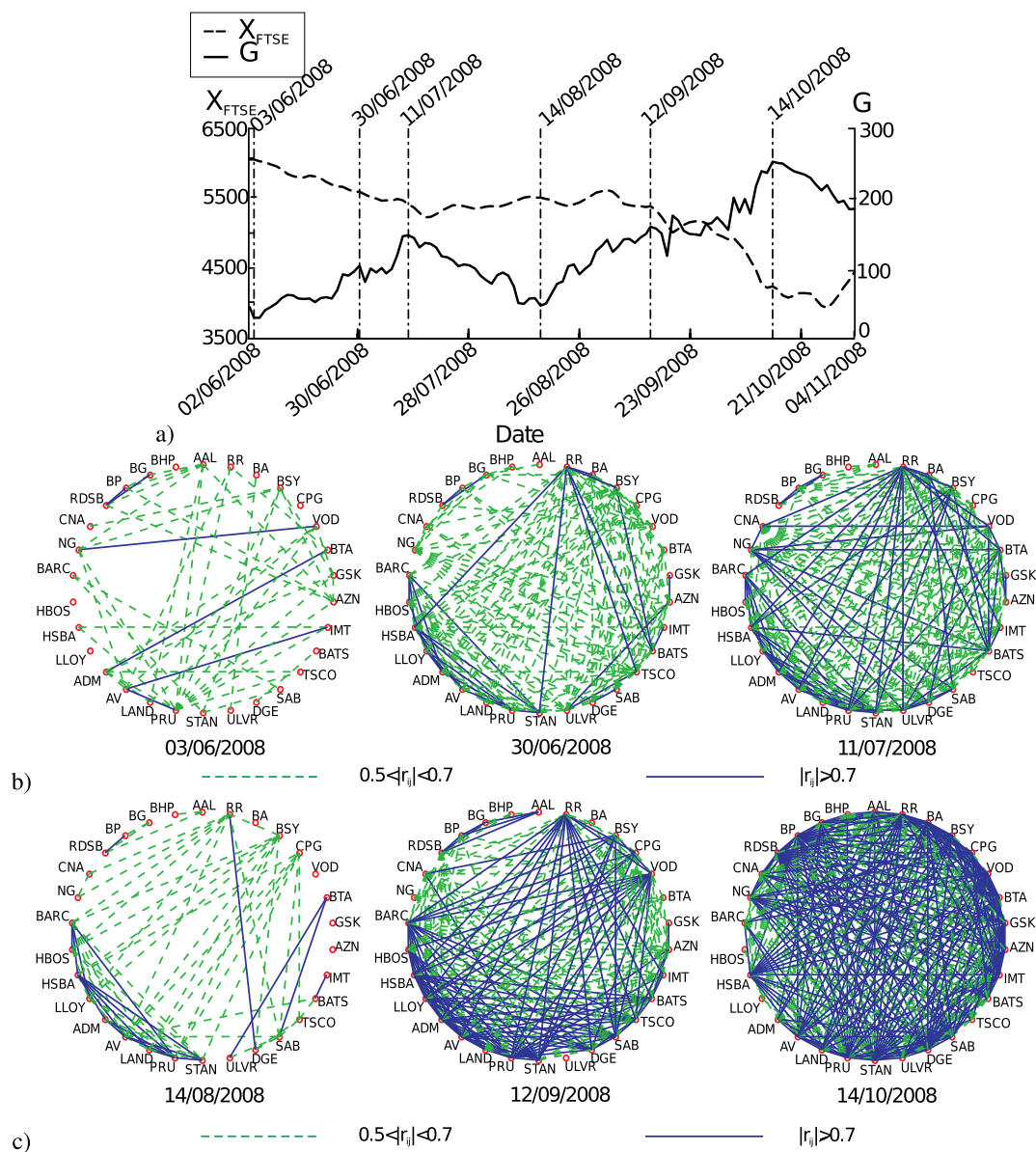


Fig. 13. Correlation graphs for six positions of the sliding time window on interval 02/06/2008 - 04/11/2008. a) Dynamics of FTSE100 (X_{FTSE} , dashed line) and of $G = \sum_{j>i, |r_{ij}|>0.5} |r_{ij}|$ (solid line), vertical lines correspond to the points that were used for the correlation graphs. b) The correlation graphs for 03/06, 30/06 and 11/07, 2008: FTSE100 monotonically decreased, and the correlation graph monotonically became more connective. c) The correlation graphs for the last three points. There was a plateau in the FTSE100 dynamics between 11/07 and 14/08 and the correlation graph on 14/08 became more rarefied. Then the crisis escalated and the connectivity grew stronger.

3. Selye's thermodynamics of adaptation

Hans Selye did not study the correlation networks. Nevertheless, his experiments and ideas shed light on dynamics of adaptation and on the network effects in adaptation. He discovered the General Adaptation Syndrome (GAS). GAS was described as the universal answer of an organism to every harmful factor or 'noxious agent'. Discovery of GAS focused several research programs on unspecific universal reactions. Analysis of Selye's experiments led him to introduce a general model of adaptation as a redistribution of the available adaptive resource to neutralize various harmful factors [147,148]. These ideas were formalized and applied to analysis of adaptation and stress in a series of

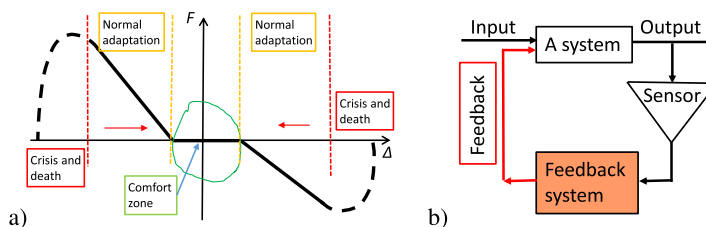


Fig. 14. The simple schemes of governors and controllers: a) Restoring force F as a function of deviation Δ of the controlled parameter from the target value. The vertical lines separate different regions of deviations: the comfort zone with the tolerance to small deviation from the target, the zone of the normal adaptation where the restoring force increases with the deviation, the crisis and death area where the possibility of adaptation are exhausted; b) The elementary scheme of the closed loop feedback control.

works [68–70,72]. In particular, increase of correlations under stress was predicted, the dynamic models of individual adaptations were created. Here we review these works with the addition of several new results. We start from the classical work of Selye’s predecessor, Walter Cannon.

3.1. Cannon’s ‘wisdom of the body’ and industrial controllers

Idea of control and feedback loop was systematically used by Cannon [23]. There are no explicit formulas in his classical book but the idea of regulation and closed loop control can be easily extracted from the homeostasis description. Two schemes in Fig. 14 illustrate these ideas. The restoring force returns the system to the vicinity of the comfort zone (Fig. 14 a). The deviation from the comfort zone can be caused by the external force. If this deviation is too large then the system cannot be returned to the normal functioning. This scheme follows Cannon’s definition: “Whenever conditions are such as to affect the organism harmfully, factors appear within the organism itself that protect or restore its disturbed balance” [23].

The feedback closed loop controller (Fig. 14 b) is an elementary brick of most controller schemes. It is well understood that the homeostasis theory, this central unifying concept of physiology, is an application of universal ideas developed in control engineering to living organisms [12,24,144]. The evolution of our understanding of homeostasis and the role of physiological regulation and dysregulation in health and disease are discussed in modern review [12]. Homeostasis is defined there as a “self-regulating process by which an organism can maintain internal stability while adjusting to changing external conditions.”

We inherited from Cannon the following picture of regulation and control of the body

- Organism is represented as a structure of relatively independent subsystems (groups of parameters);
- This decomposition is dynamic in nature and may change under the influence of harmful factors;
- Homeostasis is provided by a rich structure of feedback loops.

More and more detailed regulatory mechanisms are being disclosed and this work is supported by a data-driven approach. Small models provide qualitative understanding, while large models provide us by quantitative insights. Both mechanisms driven and data driven approaches to homeostasis modeling are developed. Among popular tools there are combinations of system analysis, dynamics, control theory, and modern data mining.

3.2. Factor – resource model of adaptation

3.2.1. Adaptation energy — generalized adaptability

Selye introduced a special currency to pay for the cost of adaptation, the Adaptation Energy (AE). He demonstrated that “during adaptation to a certain stimulus the resistance to other stimuli decreases.” His main experimental observations supported this idea showing “that rats pretreated with a certain agent will resist such doses of this agent which would be fatal for not pretreated controls. At the same time, their resistance to toxic doses of agents other than one to which they have adapted decreases below the initial value” [148].

He came to the conclusion:

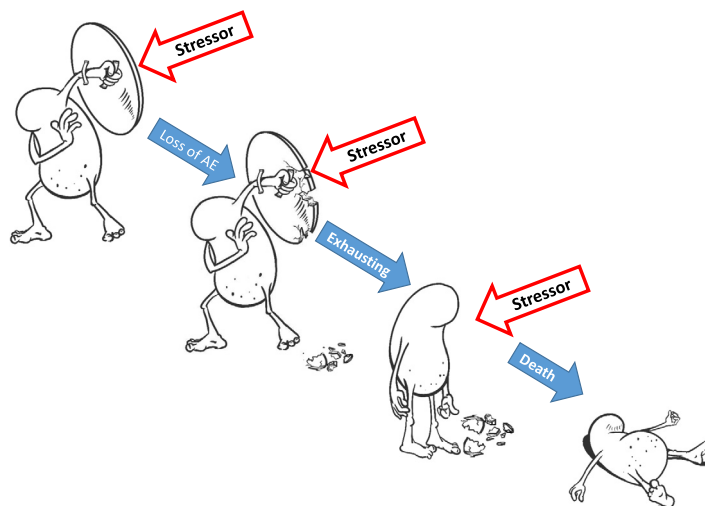


Fig. 15. Schematic representation of AE spending according to Selye's axioms. The shield of adaptation spends AE for protection from each stress. Finally, AE becomes exhausted, the animal cannot resist stress and dies. (The character 'organism' was drawn by Anna Gorban.)

- "These findings are tentatively interpreted by the assumption that the resistance of the organism to various damaging stimuli is dependent on its adaptability.
- "This adaptability is conceived to depend upon adaptation energy of which the organism possesses only a limited amount, so that if it is used for adaptation to a certain stimuli will necessarily decrease.
- "We conclude that adaptation to any stimulus, is always acquired at a cost, namely, at the cost of adaptation energy.

The idea of AE that shields the organism against stressors and is spending in this shielding is illustrated in Fig. 15.

Selye formulated 'axioms' of AE. This system of axioms was slightly edited and analyzed by his successors [143]. For modern detailed analysis we refer to [72]. The first axiom met with the most objections: *AE is a finite supply, presented at birth.*

In 1952, Goldstone published detailed analysis of Selye's AE axiomatics from the General Practitioner (GP) point of view [57,58]. From his point of view, the main problem to solve is: How one stimulus will affect an individual's power to respond to a different stimulus? Goldstone proposed the concept of a production of AE. This AE may be stored (up to a limit), as a capital reserve of adaptation. He demonstrated that this concept best explains the clinical observations and Selye's own laboratory findings. It seems to be possible that, had Selye's experimental animals been asked to spend adaptation at a rate below their AE income, they might have been able to cope successfully with their stressor indefinitely or until their AE production drops below the critical level.

These findings may be formulated as a modification of Selye's axiom 1. We call this modification *Goldstone's axiom 1'* [72]:

- AE can be created, though the income of this energy is slower in old age;
- It can also be stored as Adaptation Capital, though the storage capacity has a fixed limit.
- If an individual spends his AE faster than he creates it, he will have to draw on his capital reserve;
- When this is exhausted he dies.

Their difference from Selye's axiom 1 is illustrated by Fig. 16 (compare to Fig. 15).

The further development of Selye's axiomatics was performed by Garkavi with coauthors [55]. They developed the activation therapy, which was applied in clinic, aerospace and sport medicine.

The AE for Selye was a generalized measure of adaptability and not a physical quantity, not a type of energy. Nevertheless, the request to demonstrate the physical nature of this 'energy' was very popular. Even in the 'Encyclopedia of Stress' we can read: 'As for adaptation energy, Selye was never able to measure it...' [109]. Despite of common

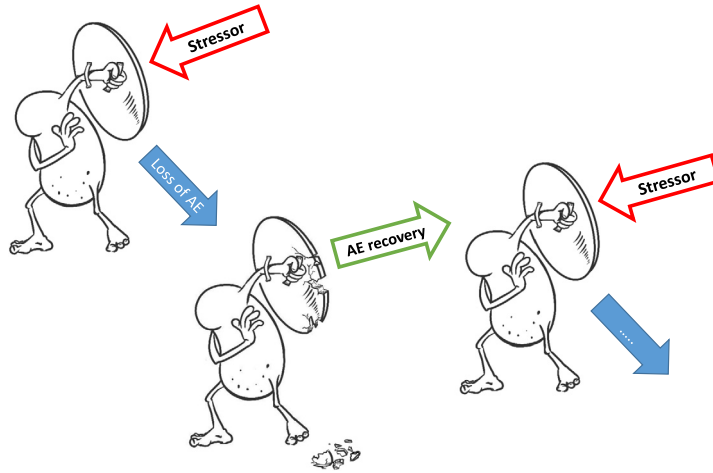


Fig. 16. Schematic representation of Goldstone's modification of Selye's axioms: AE can be restored and adaptation shield may persist if there is enough time and reserve for recovery. Surviving depends on the balance between AE loss and recovery.

use of this notion as a general 'adaptation resource' (see, for example, [17,143]) the metaphor of energy stimulated criticism of the concept: people took it literally and demanded a direct measurement of this 'energy'.

It is worth noting that the physical meaning of the free energy of adaptation has recently been revived in physical models of adaptable systems [2]. In any case, the AE can be considered as an internal coordinate on the 'dominant path' in the adaptation model, regardless of whether it has a physical interpretation or not [72].

The basic model of an adaptation resource and its use should include few details and be as simple as possible, as we expect from a model of such generality. Let us represent the life history of an organism as a sequence of adaptation events. Each event is a distribution of the available adaptation resource for neutralization of current values of harmful factors. This view reproduces the structure of Selye's experiments: a comfortable life – the action of harmful agents – a comfortable life... (See Figs. 15, 16.)

3.2.2. Factors–resource quasistatic models of adaptation

We represent the systems, which are adapting to stress, as the systems which optimize distribution of available amount of resource for the neutralization of different harmful factors (we also consider deficit of anything needful as a harmful factor).

Formally, consider systems that are under the influence of several factors $F_1, \dots, F_q$. Each factor F_i is characterized its intensity, f_i ($i = 1, \dots, q$). We accept the convention of the scale orientation: all these factors are harmful. If the fitness function W is known then fitness decreases with increasing factor intensity.

The adaptation systems are a 'shield' that protects the organism and decreases the influence of these factors. Let the organism have an available adaptation resource, R . If some amount of the resource is distributed for neutralization of factors then the effective pressure of the i th factor is $\psi_i = f_i - a_i r_i \geq 0$, where $a_i > 0$ is the coefficient of efficiency of neutralization of factor F_i and $r_i \geq 0$ is the amount of resource assigned for the neutralization of factor F_i . The condition $\sum_i r_i \leq R$ should hold. The zero value of ψ_i is optimal, and further compensation is impossible. We can take all $a_i = 1$ after rescaling, and $\psi_i = f_i - r_i$.

Thus, two quantities describe interaction of the system with a factor F_i : the factor uncompensated pressure ψ_i and the amount of resource r_i assigned for the factor neutralization.

Adaptation should optimize the allocation of resources to neutralize factors. To model such optimization, the objective function is needed. Even an assumption about the existence of an objective function and about its general properties helps in the analysis of the adaptation process. Assume that adaptation should maximize an objective function ('fitness') $W(\psi_1, \dots, \psi_q)$ which depends on the uncompensated values of factors, $\psi_i = f_i - r_i \geq 0$ for the given amount of available resource:

$$\begin{cases} W(f_1 - r_1, f_2 - r_2, \dots, f_q - r_q) \rightarrow \max ; \\ r_i \geq 0, f_i - r_i \geq 0, \sum_{i=1}^q r_i \leq R. \end{cases} \quad (4)$$

It is worth to mention that the optimal distribution of resources given by the optimization problem (4) is invariant with respect to monotonically increasing transformations of the objective functions $W \rightarrow \varphi(W)$.

The total amount of AE, R , changes in the chain of adaptation events. Selye believed that this was only consumed, but could not be produced. Goldstone proposed a more flexible concept with the body's ability to generate AE [57,58]. The first kinetic models of AE production and consumption were developed in [72].

The function W should monotonically decrease in the following sense: if $\phi_i \geq \psi_i \geq 0$ for all $i = 1, \dots, q$ then $W(\phi_1, \dots, \phi_q) \leq W(\psi_1, \dots, \psi_q)$. That is, the factors are harmful, indeed. Using this monotonicity, we can reformulate the optimality problem. Let $\sum_i f_i - R > 0$. Find $\psi_1, \dots, \psi_q$ such that

$$\left\{ \begin{array}{l} W(\psi_1, \dots, \psi_q) \rightarrow \max ; \\ f_i \geq \psi_i \geq 0, \sum_i \psi_i = \sum_i f_i - R. \end{array} \right. \quad (5)$$

If $\sum_i f_i - R \leq 0$ then all factors can be fully compensated by the allocation of resource $r_i = f_i$.

The question remains: why adaptation follows any optimality principle? Optimality is proven for microevolution processes and ecological succession. The backgrounds of ‘natural selection’ in these situations are well established after works of Haldane (1932) [76] and Gause (1934) [56]. Various concepts of *fitness* (or ‘generalized fitness’) are elaborated in many details (see, for example, review papers [14,60,98,120]). It seems productive to accept the idea of optimality and use it, while it does not contradict the data. We discuss optimality principle in the next Sec. 3.2.3.

3.2.3. The challenge of defining wellbeing and instantaneous fitness

In the optimality principle (4), (5), an individual's fitness W is used. It measures the wellbeing (or performance) of an organism. This is an instantaneous value, defined for at an instant in time. Defining of the instantaneous fitness for measuring an individual's performance is a highly non-trivial task. In mathematical biology the term ‘fitness’ is widely used in essentially another sense based on the averaging of reproduction rate over a long time [60,76,86,108,111]. This is Darwinian fitness. Darwinian fitness is non-local in time. It is the average reproduction coefficient in a series of generations and does not characterize an instantaneous state of an individual organism.

Merely highlighting the difference between individual instantaneous fitness used in the optimality principle (4), (5) and Darwinian fitness is not very useful. Optimality principles in biology have one source, the evolutionary optimality based on Darwinian fitness and on the space of possibilities where selection works (this space often deserves the epithet “mysterious space of possibilities” because of the difficulty of characterizing it). Thus, in order to use the optimality principle with individual instantaneous fitness as an objective function, we must build a bridge between this quantity and Darwinian fitness. We cannot expect this bridge to be completely constructive, but the logic of its construction should be outlined.

The synthetic evolutionary approach to optimality starts with the analysis of genetic variation and studies the phenotypic effects of that variation on physiology. Then it goes to the performance of organisms in the series of generations (with supplementary analysis of various forms of environment) and, finally, it must return to Darwinian fitness [97]. The physiological ecologists focus their attention, on the observation of variation in individual performance [134]. In this approach we have to measure the individual performance and then link it to Darwinian fitness. Empirically, Darwinian fitness is built from individual instantaneous performances, not the other way around.

The connection between individual performance and Darwinian fitness is not trivial. In particular, the dependence between them is not necessarily monotone because the individual success and success in effective reproduction do not coincide [102,129]. Fisher [49] formulated the problem of balancing individual performance and reproductive success as follows: “It would be instructive to know not only by what physiological mechanism a just apportionment is made between the nutriment devoted to the gonads and that devoted to the rest of the parental organism, but also what circumstances in the life-history and environment would render profitable the diversion of a greater or lesser share of the available resources towards reproduction.” The optimal balance between individual performance and fecundity depends on environment. This is well-known since the works of Dobzhansky [42] who stated that in the tropical zones selection typically favors lower fecundity and slower development, whereas in the temperate zones high fecundity and rapid development could increase Darwinian fitness.

Despite some controversy, the idea that the states of an organism can be linearly ordered from bad to good (well-being) is popular and useful in applied physiology. The coordinate on this scale is also called ‘fitness’. Several indicators are proposed for fitness assessment and then the fitness is evaluated as a composite of many attributes and

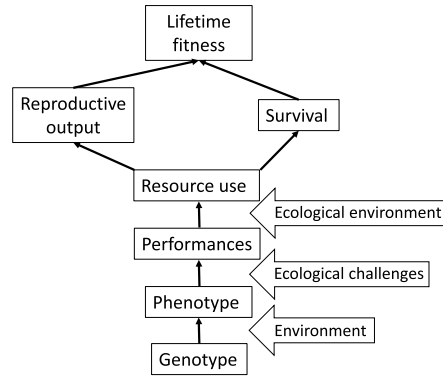


Fig. 17. The paths through from genotype to Darwinian fitness.

competencies. For example, in sport physiology these competencies include physical, physiological and psychomotor factors [138]. The combination of various components for evaluation of sport-related instantaneous individual fitness depends upon the specific sport, age, gender, individual history and even on the role of the player in the team.

In ecological physiology the evaluation of ‘performance’ is also ‘task-dependent’ [176]. It refers to an organism’s ability to carry out specific tasks (e.g., capture prey, escape predation, obtain mates). Direct instantaneous measurement of Darwinian fitness is impossible. Nevertheless, it is possible to measure various instantaneous performances several times and consider them as the components of fitness in the chain of generations. Several criteria for selection of the good measure of performance in the evolutionary study were proposed [7].

The scheme of relations between performance and lifetime fitness is sketched on flow-chart (Fig. 17) following [176] with minor changes: Genotype in combination with environment determines the organismal design (the phenotype) up to some individual variations. Phenotype determines the limits of an individual’s ability to perform day-to-day behavioral answer to main ecological challenges (performances). Performance capacity interacts with the given ecological environment and determines the resource use, which is the key internal factor determining reproductive output and survival. Darwinian fitness may be defined as the lifetime fitness averaged in a sequence of generations.

We use the instantaneous individual fitness (wellbeing) W as a characteristic of the current state of the organism, reflecting the non-optimality of its performance. The following normalization is convenient: $W = 1$ means the maximal possible performance, while $W = 0$ means inviability (death). If the organism lives at some level of W then we can consider W as a factor in the lifetime fitness. The more detailed review of the problem of defining instantaneous individual fitness is presented in [72].

For the detailed philosophical discussion of various notions of fitness we refer to [164]. It is argued that fitness is something different from both the physical traits of an organism and the number of offspring it leaves and a definition of fitness as a causal dispositional property is developed.

Finally, we have to stress that the quantitative definition of W is related to its place in the equations [72]. The change of the basic equation will cause the change of the quantitative definition. Moreover, it is plausible that for different purposes we may need different definitions of W .

3.3. Liebig’s law of the minimum and optimal resource allocation

The optimality principle (5) in its general form without specifying the fitness function $W(\psi_1, \dots, \psi_q)$ does not add much light to understanding the adaptation process. The situation changes if we add plausible hypotheses about the interaction of factors.

The fruitful idea is limitation: the worst factor determines the fitness value. That is,

$$W(\psi_1, \dots, \psi_q) = w \left(\max_{i=1, \dots, q} \{\psi_1, \dots, \psi_q\} \right), \quad (6)$$

where w is a monotonically decreasing function of one variable.

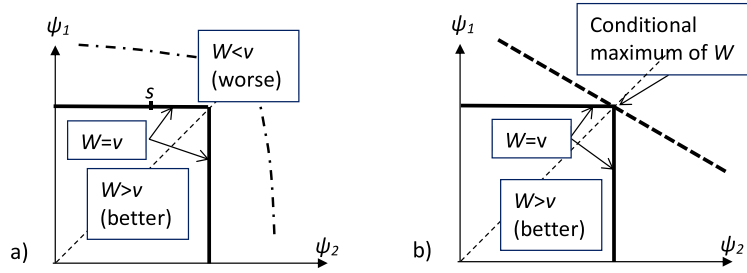


Fig. 18. (a) The law of the minimum. Coordinates ψ_1, ψ_2 are normalized values of non-compensated factors. For a given state, $s = (\psi_1(s), \psi_2(s))$, the bold solid line $\min\{\psi_1, \psi_2\} = \min\{\psi_1(s), \psi_2(s)\}$ separates the states with better conditions (higher fitness) from the states with worse conditions (a). On this line the limiting factor has the same value. The dot dash line shows the border of survival. On the dashed line (the diagonal) the factors are equally important ($\psi_1 = \psi_2$). (b) Conditional maximization of W with linear constraints. The dashed line represents the plane of linear constraints with positive normal. The conditional maximum of W on this plane is achieved on the diagonal, where the pressure of all non-compensated factors is the same.

This form (6) of the *law of the minimum* assumes correct normalization of factors and evaluation of the pressure of factors in the same units of measurement.

The function $w(\max_{i=1,\dots,q}\{\psi_1, \dots, \psi_q\})$ is not fully defined because it includes unknown monotonically decreasing function w of one variable. Nevertheless, the level sets do not depend on the specific w . The superlevel set given by the inequality $W > v$ for a value v is a cube (see Fig. 18). Maximum of W in the positive orthant on a hyperplane with positive normal belongs to the diagonal, where all the factor pressures are equal (Fig. 18 b). The optimality problem (5) for functions (6) is a bit more complex and requires to maximize W in the intersection of the parallelepiped $f_i \geq \psi_i \geq 0$ with the hyperplane L given by the equation $\sum_i \psi_i = \sum_i f_i - R$. Nevertheless, it can be solved in explicit form. For any objective function W that satisfies the law of the minimum in the form (6) the optimizers r_i in the problem (4), (5) for functions (6) are defined by the following algorithm.

1. Re-enumerate factors in the order of their intensities starting from the worst: $f_1 \geq f_2 \geq \dots f_q$.
2. Calculate differences $\Delta_j = f_j - f_{j+1}$ (take formally $\Delta_0 = \Delta_{q+1} = 0$).
3. Find such k ($0 \leq k \leq q$) that

$$\sum_{j=1}^k j \Delta_j \leq R \leq \sum_{j=1}^{k+1} j \Delta_j.$$

For $R < \Delta_1$ we put $k = 0$, for $R > \sum_{j=1}^{k+1} j \Delta_j$ we take $k = q$.

4. If $k < q$ then the optimal amount of resource r_{j_l} is

$$r_l = \begin{cases} \Delta_l + \frac{1}{k} \left(R - \sum_{j=1}^k j \Delta_j \right), & \text{if } l \leq k+1; \\ 0, & \text{if } l > k+1. \end{cases} \quad (7)$$

If $k = q$ then $r_i = f_i$ for all i .

After all we have to restore the initial enumeration $f_{i_1} \geq f_{i_2} \geq \dots f_{i_q}$. This optimization is illustrated in Fig. 19.

We demonstrated here the *law of the minimum paradox* [69]: if for a randomly chosen pair “organism–environment” the law of the minimum typically holds (6), then, in a well-adapted system, we have to expect violations of this law and the pressure of factors tend to be less different (see Fig. 19). In particular, the factors with the strongest pressure can become equally important (factors f_{1-4} in Fig. 19).

This effect also provides a mechanism of increase of correlations and variance under strong load of an external harmful factor. For strong limitation and strong load of the limiting factor the main difference between organisms is the uncompensated pressure of the limiting factor (Fig. 20a) and near colimitation different factors can become limiting ‘by chance’ and the organisms deviate from the comfort zone in different directions (Fig. 20b). According to the general results about quasiorthogonality of random vectors in high-dimensional spaces [71], these directions are almost orthogonal (with high probability) and the *dimension of the data cloud can be effectively estimated as the number of independent co-limiting factors*.

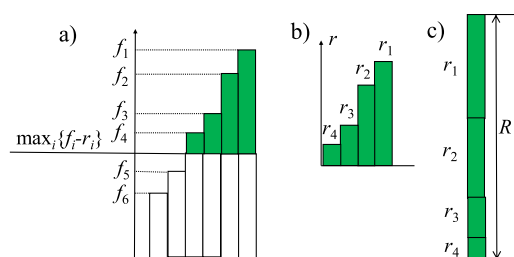


Fig. 19. Optimal distribution of resource for neutralization of factors under the law of the Minimum. (a) Histogram of factors intensity (the compensated parts of factors are highlighted, $k = 3$), the distribution of compensated factors became more uniform than the initial distribution of factor loads, (b) a histogram of resource allocation, (c) the sum of allocated resources.

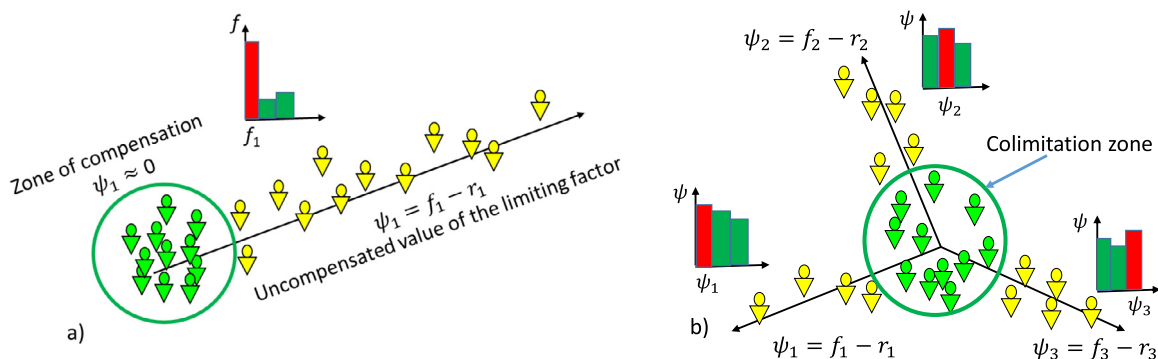


Fig. 20. High correlations in the situation with strong limitation and uncompensated limiting factor (a) and low correlations near colimitation, where factor pressures are almost equal (b).

The observation that adaptation makes the limiting factors equally important was supported by many data of human adaptation to the Far North conditions [68]. At the same time, disadaptation causes inequality of factors and leads to appearance of single limiting factor. The results are used for monitoring of human populations in Far North [146].

The “law of the minimum” in its original form stated that growth is controlled by the scarcest resource (limiting factor) [140]. This law was published by Justus von Liebig’s in 1840 and is very often called “Liebig’s law”. Twelve years earlier, Carl Sprengel published an article that contained in essence the law of the minimum. Thus, it can be called the Sprengel–Liebig law of the minimum [168]. The history, of this law, its applications and related controversies were reviewed in [69].

The law of the minimum is one of the most important tools for mathematical modeling of ecological systems. This provides the clue to creating the first model of multi-component and multi-factor systems. This clue sounds pretty simple: first of all, we must take into account the most important factors, which are probably limiting factors. Everything else should be eliminated and returned back only if ‘sufficient reason’ is found. This law was considered as an “archetype” for ecological modeling [119].

The economical metaphor of ecological concurrency was elaborated for combination of the idea of limitation with optimization approach [13,29]. In this approach, a plant should adjust allocation so that, for a given expenditure in acquiring each resource, it achieves the same growth response: Growth is equally limited by all resources. This is a result of adjustment: adaptation makes the limiting factors equally important.

Three inter-related issues which form the core of evolutionary ecology were considered in [152]: (i) key environmental factors, (ii) organismal traits that are responses to the key factors, and (iii) the evolution of these key traits. Adaptation leads to optimality and equality of traits as well as of factors but under variations some traits should be more limiting than others, therefore the concept of ‘limiting traits’ was introduced [152].

Analysis of species-specific growth and mortality of juvenile trees at several contrasting sites suggests that light and other resources can be simultaneously limiting, and challenges the application of the Law of the Minimum to tree sapling growth [89]. This is the manifestation of the “colimitation” phenomenon: limitation of growth and surviving by a group of equally important factors and traits.

There are regions in the world ocean where chlorophyll concentrations are lower than expected concentrations given the phosphate and nitrate levels. Limitations of phytoplankton growth by other nutrients like silicate or iron was studied and supported by experiments [8].

The double–nutrient–limited growth appears also as a transition regime between two regimes with single limiting factor. This phenomenon was studied for bacteria and yeasts at a constant dilution rate in the chemostat, and three distinct growth regimes were recognized: a carbon-limited regime, a nitrogen-limited regime, and the double–nutrient–limited growth regime where both the carbon and the nitrogen source were below the detection limit. The double–nutrient–limited zone is very narrow at high growth rates and becomes broader during slow growth [46,183].

There is colimitation on community level even if organisms and populations remain limited by single factors. Communities adjust their stoichiometry by competitive exclusion and coexistence mechanisms. This conclusion was supported by a model and experiments with microcosms [36]. Various types of colimitation were identified and studied and the potential nutrient colimitation pairs in the marine environment were described in [141].

The multiple-resource limitation responses in plant communities was summarized using datasets from 641 studies of nitrogen (N) and phosphorus (P) in freshwater, marine and terrestrial systems. Synergistic response to N and P addition was found in more than half of the studies [77].

The ‘trade-off’ between law of the minimum and co-limitation includes optimization arguments and becomes paradoxical [69,157]: if organisms in a randomly selected environment satisfy the law of the minimum (everything is limited by the factor with the worst value) then, after adaptation, these factors tend to be equally important. Conversely, if for an organism in a randomly selected environment importance of factors is comparable and factors superlinearly amplify each other then, in the process of adaptation, such systems evolve towards the law of the minimum and less factors become important.

The main consequences of the law of the minimum for the topic of our paper are:

- In well-adapted systems the most harmful external factors are approximately equally important. The behavior of the system depends on fluctuations of many factors, and therefore, the data cloud is expected to be multidimensional. This conclusion applies both to data on populations (groups) of similar well-adapted systems and to data on different snapshots from the same system.
- Conversely, when the pressure from a factor suddenly increases, then the behavior of systems mainly depends on the uncompensated pressure of this factor. The expected data cloud has a lower dimension (with a large dominance of this factor). It can even become almost one-dimensional with accumulation of almost all variance of the relevant attributes in the first principal component.
- The number of independent co-limiting factors gives the effective estimate of the dimension of the data cloud.

3.4. Generalized law of the minimum

In fact, when the pressure of several factors is comparable, then we can expect that all of them affect the situation, and the law of the minimum has its original form asymptotically, when the pressure of one factor is much greater than the others. The levels of the fitness functions (Fig. 18) should be smoothed near the bisector of the positive octant, where the pressures of factors are close to each other (Fig. 21). In this area we expect colimitation.

Fig. 21 illustrates the *generalized law of the minimum*: when the pressure of the factors differs significantly, then fitness is determined by the worst factor, but for closer values of pressure several factors matter because fitness is a smooth function of the pressure of the factors. This qualitative idea needs proper formalization. Notice that in both Figs. 18 and 21 the set where $W > v$ is convex. This property that the superlevel set is convex may be crucial. Compare: in thermodynamic convexity/concavity of thermodynamic potentials is a very seminal idea.

We know that the superlevel sets of concave functions are concave. But concavity of functions requires more: a function defined on a convex set U is convex if and only if the region below its graph is a convex set. This definition involves both the arguments of the function and its values: $W(\mathbf{x})$ is concave if the set of pairs $(\mathbf{x}, v)$ is convex, where $v \leq W(\mathbf{x})$ and $\mathbf{x} \in U$.

In physics, the value of energy, entropy, free energy and other thermodynamic potentials play important role. In the principle of fitness maximization, the values of fitness are not much important, but the order of these values matters. It measures which combination of traits is ‘better’. All the solutions of the fitness maximization problem will remain the same after a transformation of the scale $W \rightarrow \vartheta(W)$ with monotonically increasing function ϑ . Concavity

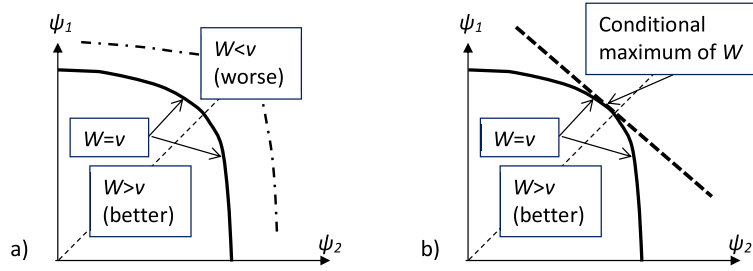


Fig. 21. (a) The generalized law of the minimum. Coordinates ψ_1, ψ_2 are normalized values of non-compensated factors. The fitness function W is quasiconcave that means convexity of the superlevel sets ($W \geq v$). The dot dash line shows the border of survival. The bold solid line is the level set of fitness, $W = v$. It separates the states with better conditions (higher fitness) from the states with worse conditions. On the dashed line (the diagonal) the factors are equally important ($\psi_1 = \psi_2$). (b) Conditional maximization of W with linear constraints. The dashed line represents the plane of linear constraints L with positive normal. The conditional maximum of W on this plane is achieved at an extreme point of the convex superlevel set of W where L is a support hyperplane.

of the function is not invariant with respect to this transformation, whereas its quasiconcavity persists. A function $W(\mathbf{x})$ defined on a convex set U is quasiconcave if and only if its superlevel sets given by inequalities $W(\mathbf{x}) > v$ are convex. The equivalent definition is: $W(\mathbf{x})$ defined on a convex set U is quasiconcave if for all $\mathbf{x}, \mathbf{y} \in U$ and $\lambda \in [0, 1]$ inequality holds

$$f(\lambda x + (1 - \lambda)y) \geq \min \{f(x), f(y)\}.$$

The generalized law of the minimum for the system of harmful factors is defined as follows: the fitness $W(\psi_1, \dots, \psi_q)$ defined in the non-negative orthant ($\psi_i \geq 0$) has three following properties:

1. Monotonicity. If $\phi_i \geq \psi_i \geq 0$ for all $i = 1, \dots, q$ then $W(\phi_1, \dots, \phi_q) \leq W(\psi_1, \dots, \psi_q)$.
2. Quasiconcavity. The superlevel sets defined by the inequalities $W(\psi_1, \dots, \psi_q) \geq v$ are convex.
3. For any i , $\frac{\partial W(\psi_1, \dots, \psi_q)}{\partial \psi_i} \rightarrow 0$ when $\psi_i \rightarrow 0$, other values of ψ_j ($j \neq i$) are constant, $\psi_j \geq 0$ and at least one $\psi_j > 0$.

An additional requirement #3 is needed to ensure the asymptotic validity of the law of the minimum for a situation where the pressure of factors differs significantly: the effect of the factors with small pressure vanishes.

Such a function is presented in Fig. 21a). The conditional maximum of W on a hyperplane L with positive normal is achieved at an extreme point of the convex superlevel set of W where L is a support hyperplane (Fig. 21b)). We can call the system of factors that satisfy the generalized law of the minimum the *generalized Liebig system*.

Typical examples of generalized Liebig systems give the fitness functions defined through the distance of the vector of uncompensated factor loads $\boldsymbol{\psi} = (\psi_1, \dots, \psi_q)$ from the origin that is the state with fully compensated loads:

$$W(\boldsymbol{\psi}) = \begin{cases} 1 - \frac{\|\boldsymbol{\psi}\|_p}{\psi_0}, & \text{if } \|\boldsymbol{\psi}\|_p \leq \psi_0; \\ 0, & \text{if } \|\boldsymbol{\psi}\|_p > \psi_0. \end{cases} \quad (8)$$

Here, $\|\boldsymbol{\psi}\|_p$ for $p \geq 1$ is the l_p norm:

$$\|\boldsymbol{\psi}\|_p = \left(\sum_i |\psi_i|^p \right)^{1/p}. \quad (9)$$

Unit circles of l_p norms are presented in Fig. 22. For $0 < p < 1$, formula (9) does not define a norm because the ‘ball’ becomes not convex. Such uniform functionals without the triangle inequality are called ‘quasinorms’.

For $p > 1$, the fitness (8) has the properties 1-3. For $p = 1$, the property 3 is not satisfied because in this case the partial derivatives of W do not depend on (positive values of ψ_i). The classical Liebig law corresponds to the case $p = \infty$ because $\|\boldsymbol{\psi}\|_\infty = \max_i \{|\psi_i|\}$.

This model fitness (8) $W = 1$ when the factor pressures are compensated, $\boldsymbol{\psi} = 0$, and goes to zero when $\|\boldsymbol{\psi}\|_p$ becomes larger than a threshold $\psi_0 > 0$. Zero fitness means death.

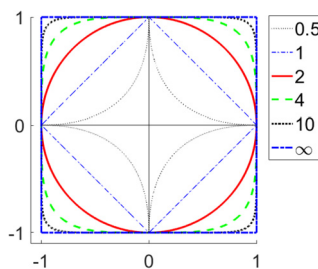


Fig. 22. Unit circles for L_p norms with $p = 1, 2, 4, 10, \infty$ and for a quasinorm with $p = 1/2$.

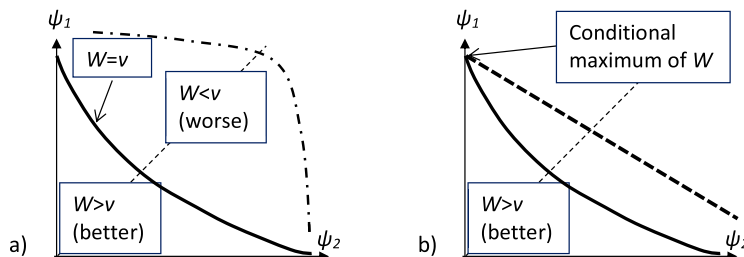


Fig. 23. (a) Synergistic system of harmful factors. Coordinates ψ_1, ψ_2 are normalized values of non-compensated factors. The fitness function W is quasiconvex that means convexity of the sublevel sets ($W \leq v$). The dot dash line shows the border of survival. The bold solid line is the level set of fitness, $W = v$. It separates the states with better conditions (higher fitness) from the states with worse conditions. On the dashed line (the diagonal) the factors are equally important ($\psi_1 = \psi_2$). (b) Conditional maximization of W with linear constraints. The dashed line represents the plane of linear constraints L with positive normal. The conditional maximum of W on this plane in the positive orthant is achieved at the boundary of the orthant where some factors are completely compensated.

3.5. Synergistic systems of factors

The law of the minimum means that everything is defined by the worse factor. The generalized law of the minimum gives qualitatively similar but ‘smoothed’ picture (see Fig. 21). Importance of the Liebig principle and its consequences immediately leads us for the question: how the strong deviations of the Liebig principles look like and what are their consequences for adaptation in the frame of the Selye’s thermodynamic approach given by the optimality principle (5).

According to the generalized law of the minimum, fitness is a quasiconcave function and its sensitivity to minor factors vanishes. The opposite situation is also possible, when factors amplify the harm from other factors.

The system of harmful factors is called *synergistic* if the fitness $W(\psi_1, \dots, \psi_q)$ defined in the non-negative orthant ($\psi_i \geq 0$) has two following properties:

1. Monotonicity. If $\phi_i \geq \psi_i \geq 0$ for all $i = 1, \dots, q$ then $W(\phi_1, \dots, \phi_q) \leq W(\psi_1, \dots, \psi_q)$.
2. Quasiconvexity. The sublevel sets defined by the inequalities $W(\psi_1, \dots, \psi_q) \leq v$ are convex.

An illustrative example of synergistic system of factors is presented in Fig. 23.

Formula (8) presents a synergistic system when $0 < p < 1$. The case $p = 1$ is the bordercase: both sublevel and superlevel sets of W are convex, and sensitivity to all factors is the same constant $1/\psi_0$ (when $W > 0$). For $0 < p < 1$, the sensitivity $\frac{\partial W}{\partial \psi_i}$ is larger for smaller ψ_i and increases to infinity when $\psi_i \rightarrow 0$ (and when other p_i are constant). Such a high sensitivity to small concentrations of some chemicals in combination with other chemicals has been widely discussed.

For a synergistic system, adaptation tends to compensate some factors completely, and the dimensionality of the data cloud is expected to decrease with adaptation. Let us demonstrate this on the basic example – the fitness (8) with $0 < p < 1$. A simple calculation shows that for a given vector of factor loads, $\mathbf{f} = (f_1, \dots, f_q)$ and the available total resource R the optimal distribution of the resource can be constructed by simple iterations. Let (after reordering) $f_1 > \dots > f_q$.

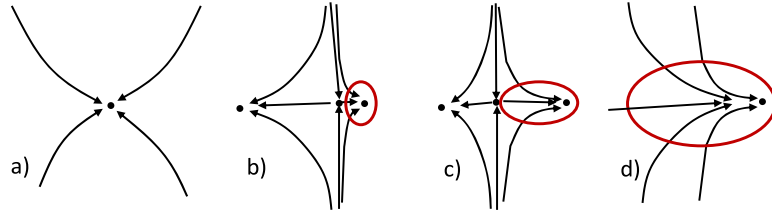


Fig. 24. Schematic representation of the bifurcation model: a) the globally stable healthy state; b) after the first bifurcation – the disease state appeared (encircled by red); c) disease expansion – attraction area of the healthy state decreases; d) after the second bifurcation – the healthy state disappeared.

1. If $R < f_q$ then compensate f_q as much as it is possible: set $r_q = R$, $\psi_q = f_q - R$ and stop.
2. If $R \geq f_q$ then compensate f_q completely, set $r_q = f_q$, $\psi_q = 0$.
3. Set $R := R - f_q$ and iterate with $q - 1$ factors $f_1 > \dots > f_{q-1}$.

Optimization for general synergistic systems was analyzed in [69]. There is significant difference between adaptation to the synergistic system of factors and to the generalized Liebig system: in synergistic systems resource is distributed for complete neutralization of small factor pressures with the highest sensitivity, and the dimensionality of data tends to decrease, while for a generalized Liebig system, adaptation tends to reduce the strongest differences between factor pressures and the dimensionality of data tends to increase in the course of adaptation.

Optimization principle is a basis of the analysis of trade-offs between various functional traits under pressure of various environmental factors. The effect of shapes of the objective functions was clearly demonstrated by Tilman for ecological problems [162,163]. A general approach to evaluate the effect of different trade-off shapes on species co-existence based on the modeling of invasion experiments is presented in [47]. Application of the general optimization principle (4), (5) to physiological adaptation was proposed in [68], and developed further in [69,70,72,146].

Qualitatively, the main difference is usually considered between convex and concave functions (or ‘trade-off curves’ when analyzing pairwise relations). It is obvious that the concavity and convexity can be substituted by the quasiconcavity and quasiconvexity without the change of the results. These properties are invariant with respect to monotonically increasing transformations of the scale of fitness. There are many other, combining versions of the trade-offs. For example, there can be groups of synergistic factors with ‘Liebig’s’ relations between these groups.

3.6. Alternative theories of low dimension and high correlations under stress

3.6.1. Bifurcations, slow manifolds, critical retardations and critical fluctuations

Some alternative theories suggest that observed increase in correlations is the result of bifurcation in the dynamical system of physiological regulation [32]). This dynamical system is hidden behind the observed dynamics of physiological parameters. Before this bifurcation, a regime of normal functioning is stable in wide area, that is stable homeostasis. The hypothetical bifurcation leads to appearance of a new attractor, a locally stable regime of illness. As the simplest version, we can consider appearance new attractor as a saddle-node bifurcation [92] (see Fig. 24). With development of disease the basin of attraction of the normal homeostasis contracts and, finally, vanishes in the second saddle-node bifurcation. This evolution of phase portrait is illustrated in Fig. 24, from Fig. 24a to Fig. 24d.

The increase of correlations near bifurcations means that there appear “slow modes” or, the dynamics (after some initial layer) approaches slow manifold (in the simplest cases this manifold is a curve), and the systems near bifurcation are distributed along this low-dimensional manifold. The state space of hidden dynamics is not fully observed, but an increase in correlations that follows a decrease in effective dimensions near bifurcations can be observed in almost any set of relevant variables.

The approximately slow manifold in the illustration (Fig. 24) is represented by the trajectories that lead from the unstable steady state (a saddle) to the stable nodes. After some initial period, the motion approaches this slow manifold (a curve). One-dimensional motion along this curve can always be represented as a gradient system in a potential well with ‘energy’ $V(x)$, where x is the coordinate along the curve:

$$\frac{dx}{dt} = -\frac{dV}{dx}.$$

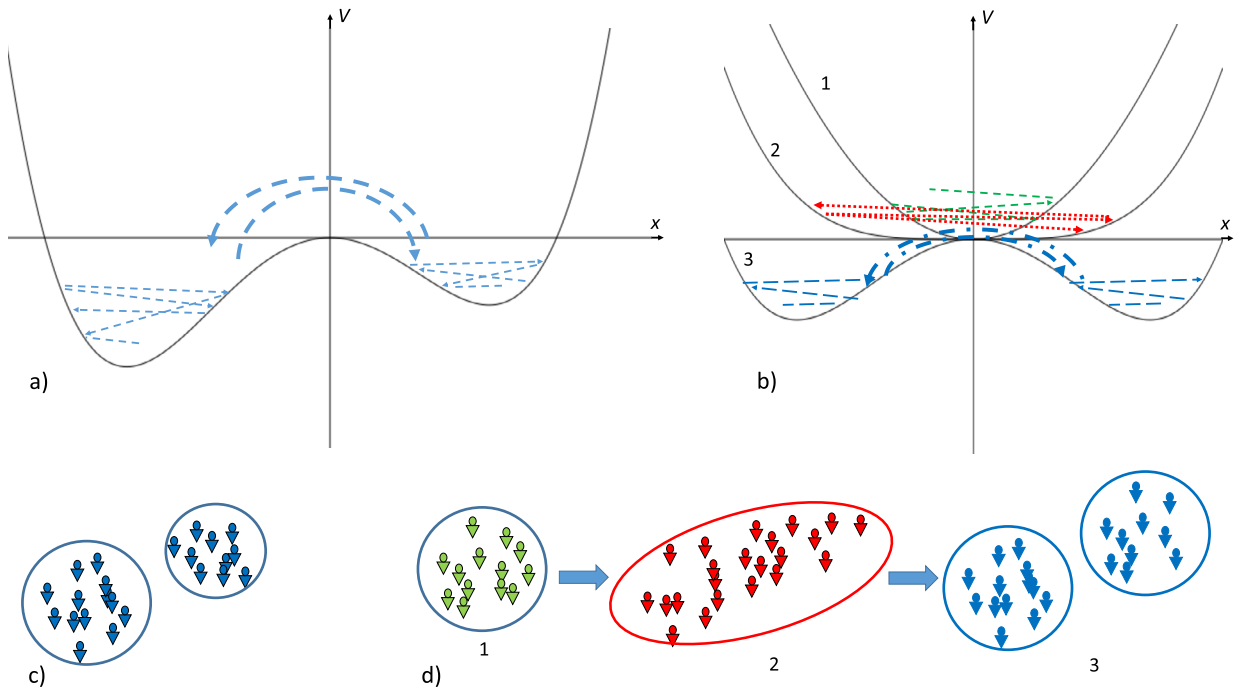


Fig. 25. Fluctuations in the bifurcation model. a) The potential wells for two stable and one unstable states. The thin dashed lines symbolize fluctuations in the potential wells near the stable states. The bold dashed lines illustrate jumps between potential wells due to large fluctuations. b) Disintegration of a stable equilibrium into three steady states: two stable and one unstable. Configurations of potential wells for three stages are presented: (1) one stable state with non-degenerated quadratic term of the potential, (2) degeneration of the quadratic term, and (3) two potential wells for two symmetric stable steady states. Fluctuations are presented by dashed, dotted and dash-dotted lines. c) Expected distribution of population in the potential wells of two steady states from the panel a); d) Evolution of the distribution of population for the evolution of the potential wells presented in panel b).

This representation is illustrated in Fig. 25a.

Various fluctuations in the environment are imminent life conditions. There are two types of fluctuations in Fig. 25a: jiggling around the equilibria and large jumps between different basins of attraction. Such jiggling and jumps lead to a clustered stationary distribution (Fig. 25c).

Increase of variance is a particular case of the famous effect of increase of fluctuations near bifurcations and phase transitions [45]. The relaxation to the attractor may have anomalously long critical delays [61,62,64] and, therefore, the random change of environment can add large fluctuations. A brief digest of these works was published recently [63].

The essence of the approach based on bifurcations, critical slowing down and critical fluctuations, can be formulated in the following way: *nothing special, just the dynamics*. The qualitative picture behind this theory is quite simple. Organism is a large-dimensional dynamical system that depends on many parameters. Parameters change in time due to many reasons (environment, illness, aging, etc.). Attributes are functions on the phase space. When the system approaches bifurcations, the slow modes appear, the effective dimensionality decreases and the variance increases. It could be worth to mention that this point of view on the bifurcations and slow modes was developed first by L. Landau in his work about turbulence [95]. The logic was very similar: the system is very multidimensional, and we do not know the details, but when the laminar flow loses stability and a new attracting (periodic) regime is born, then a slow mode appears, and we can study its dynamics separately, without consideration of the entire system. This slow mode can be defined as a measure of deviation from the laminar flow.

The contradictions between this theory and the AE approach are not obvious: both consider low-dimensional dynamics under stress as a mechanism for increasing correlation. Near bifurcations, the AE can be considered as the coordinate along the slow manifold. Nevertheless, there exist some differences in details because the AE approach depends on the organization of the system of factors and gives different results for the Liebig and the anti-Liebig

(synergistic) systems of factors. The interpretation of the approach to death is also unclear in the hidden bifurcation methods, whereas interpretation of death through exhaustion of the adaptation resource was introduced by Selye and supported by his experiments.

A verifiable difference is in the details of the correlation increase. If after bifurcation two attractors appear at non-zero distance (see Fig. 24 and Fig. 25a), then the cloud of data is expected to have a clustered structure (Fig. 25c). The apparent increase in correlations in this case just reflects the difference between attributes in clusters. For systems with strong limitation such clusters are not expected (Fig. 20a).

One important comment should be added for multidimensional systems: bifurcation has a significant effect only on a part of the variables. This is the relevant dominant group, and one of the first goals of the researcher is to identify these variables [32].

Revealing of the relevant dominant group of attributes is important in a much broader context. When studying any phenomena in a multidimensional system, we must identify the variables that are directly involved in this process. These variables can form a special group of already known variables or must be constructed from them. In physics of phase transitions, these are the classical ‘order parameters’. In the theory of bifurcations, the projection of a system on the relevant normal form selects the analog of the order parameter (see, for example, [92]). Selection of the relevant dominant groups is always needed in the adaptation research. Sometimes, it is performed explicitly, with relation to the specific theory about the nature of the dimensionality reduction (self-simplification or degeneration) [32] or, often, without specification of such a theory (for example, [25–27,75]). Most often, the selection of relevant attributes is performed implicitly as a prerequisite for research. For example, the choice of lipid fractions for a comparative analysis of metabolism in healthy newborns in the temperate zone of Siberia and a Far North city was not accidental [68] (see Fig. 1). At the beginning of this study, a biologically plausible hypothesis about the relevance of these variables was accepted.

Sometimes the assumption about bifurcations (tipping points) can be supported by the data [32], but often, identifying hidden bifurcations is a non-trivial task, and they are postulated rather than demonstrated. In that sense, the hidden AE is substituted by another assumption, hypothetical bifurcations. However, explaining Selye’s experiments [147,148] in terms of bifurcation remains an open challenge.

It cannot be ruled out that the dominant path of the organism’s dynamics, parametrized by the adaptive energy, gives one more description of the hidden bifurcations, but such a theory still needs to be developed.

3.6.2. Adaptation dynamics near a border of surviving

We used the idea of adaptation resource to demonstrate how can the models of adaptation explain the observable dynamics of correlation and variance under environmental loads. Another theoretical approach to explain these effects was proposed in [135] and developed further with applications in [136,150]. The idea is quite simple. Each organism is presented by a point in the n -dimensional space $\mathbb{R}^q$ of physiological attributes. The ‘area of existence’ is a bounded set $U \subset \mathbb{R}^q$ with smooth boundary. The population is a cloud of points in U . At this scale, individual differences in properties and life conditions can be considered as random. The individual history can be represented by a random walk inside U . At the population level, this movement can be modeled as diffusion. In the absence of the pressure strong common factors diffusion distributes population in area of existence (Fig. 26a).

The pressure of a common factor that acts similarly on all the organisms in the population can be modeled as an external ‘wind’ (advection) (Fig. 26b). It presses the population to the border of U as it is represented in (Fig. 26b,c). if intensity of this external pressure increases then the population concentrates in close vicinity of the boundary (Fig. 26c). It looks as a disk near the border. The detailed analysis shows that for large intensity of the factor, f , the thickness of this disk goes to zero as $1/f$, whereas its diameter decays slower, as $1/\sqrt{f}$ and it becomes essentially $n - 1$ -dimensional cloud. For formal detail we refer to Appendix B.

To check whether this mechanism is close to reality we have to resolve the “disk–or–bullet” alternative: How does the correlation increase: (i) the largest eigenvalues of the correlation matrix increase and the ellipsoid of concentration becomes stretched-out in small amount of dimensions (presumably, in one, that is the bullet form) or (ii) the smallest eigenvalue goes fast to zero, and the ellipsoid of concentration becomes flattened out (the disk or the pancake form). The existing data ensure us that this alternative should be resolved in favor of bullet: under the pressure of environment factors, the small amount of largest eigenvalues increases. It should be noted that the variance in this model decreases with increasing pressure of the factor. That also contradicts the majority of existing observation.

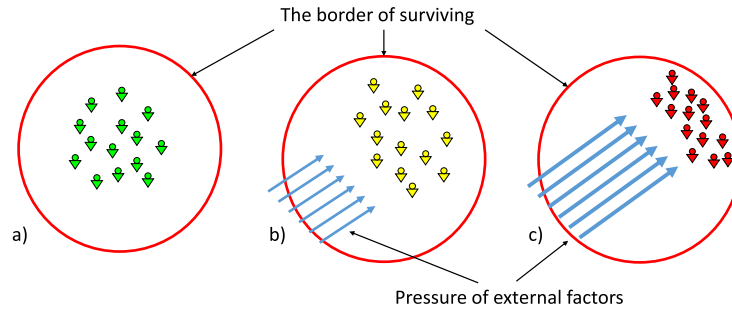


Fig. 26. Under the pressure of external factors the population approached the border of existence and becomes effectively a cloud concentrated near a hyper-surface (one dimension less).

4. Selye–Goldstone kinetics: dynamics of individual adaptation

4.1. The type of model we need

Quasi-static modeling of adaptation as a chain of optimal allocation of resources to neutralize the pressure of harmful factors is simple, allows for a detailed analysis of models and provides a tool for analyzing adaptation experiments that are organized in the form of such a chain of adaptation challenges and responses to them. Nevertheless, many problems remain open. For example: How long lasts the protection from a harmful factor provided by the allocated amount of AE? How much AE will be available for distribution at the next time step? The answers are impossible in a such reduced formalism because it does not include time. A system of dynamic models is needed. It is a shift from a view inspired by classical thermodynamics to the kinetics of production and expenditure of AE.

The status of these models should be clearly represented. There are two basic approaches to modeling, *bottom-up* and *top-down*, and the emerging *middle-out* approach. The bottom-up approach starts from a detailed picture and many additional experimental studies are usually required for identifying the details of the elementary events [54]. The top-down approach to modeling is closer to classical physiology [79]. It starts from very general integrative properties and then adds some details. To combine the advantages of the bottom-up and the top-down approaches, the *middle-out* approach was used by many authors and analyzed specially in several works [15,90]. With this approach, modeling starts neither from the top level, nor from the bottom, but from the level which is ready for formalization, where one can plausibly guess the main mechanism.

We use and elaborate further in this section Selye’s empirical insights about the forms of AE, follow the middle-out approach to modeling, and exploit the kinetic formalisms of mathematical chemistry and ecology.

4.2. Kinetic factor–resource models

In most experiments and in many clinical situations, a dominant harmful factor can be identified. In particular, in Selye’s experiments [147,148], rats were exposed to one noxious agent each time (but can be exposed to different agents sequentially). The dynamic model presented in this Subsection describes adaptation to one factor with intensity f .

The scheme of the model (Fig. 27) completely represents the system of equations, if we define all the quantities, mention that $W(\psi/\psi_0) = 1 - \psi/\psi_0$ (if $0 \leq \psi/\psi_0 \leq 1$), and formulate the rules of opening and closing the tap between the reserve AE storage (*deep AE* in Selye’s terminology) and the main AE storage (*superficial AE*). Goldstone’s comments about AE productions for realistic models of adaptation [57,58] are taken into account in this model (see Fig. 16). The resource and reserve storages have limited capacity (R_0 and R_{rv} , correspondingly) and the production flux is proportional to the product of fitness W (the measure of wellbeing) by the deficit of AE: $R_0 - r_0$ for the main storage, and $R_{rv} - r_{rv}$ for the reserve.

Historically, the first set of ideas was formalized in a model with two variables: the stored AE r_0 and the amount of AE r allocated to neutralize the harmful factor. In Fig. 27, this part of the diagram is outlined with a dashed line. Recall that $\psi = f - r$ is the uncompensated pressure of the factor. R_0 is the maximal amount of AE in the storage. The production flux is proportional to the product of fitness W (the measure of wellbeing) by the deficit of AE in the

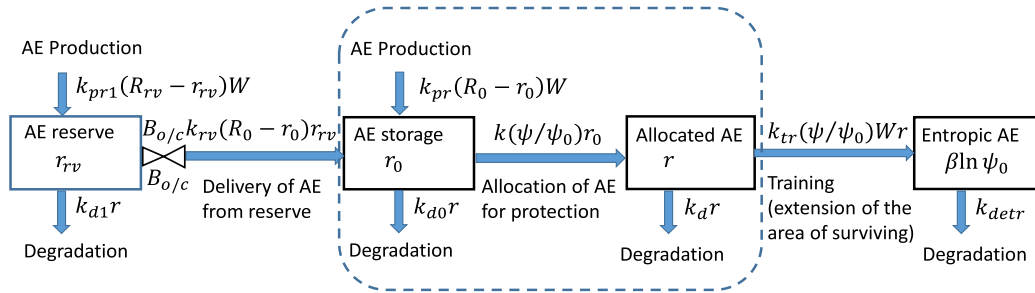


Fig. 27. The scheme of kinetic model of AE production and delivery. Four quantities describe the distribution of AE “main storage” (resource) r_0 , resource r , allocated to neutralization of the harmful factor, “reserve” r_{rv} and the “entropic form of AE” accumulated in the training results, $\beta \ln \psi_0$. AE flows are represented by arrows with the flow rate written above the arrows. The state of the tap between the reserve and the main storage is described by the Boolean variable B_0/c .

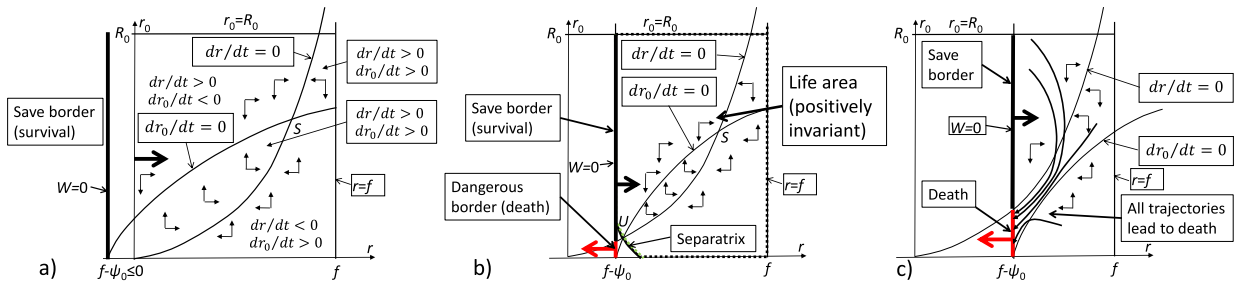


Fig. 28. Safe and dangerous borders for dynamics of adaptation in the simplest model. The r -nullcline cuts the vertical line $W = 0$ (the border of death) into two parts: $\dot{W} < 0$ (dangerous border, red) and $\dot{W} > 0$ (safe border, black). (a) If $f < \psi_0$ then the whole border is safe. The nullclines have unique intersection point S , the globally stable node. (b) If the r - and r_0 -nullclines have two intersections, the stable (S) and unstable (U) equilibria, then the separatrix of the unstable equilibrium U separates the attraction area of the dangerous border of death from the area of attraction of the stable node (life area). (d) If there exists no intersection of the nullclines then all the trajectories are attracting to the dangerous border.

storage, $R_0 - r_0$. The allocation flux is proportional to the normalized value of the uncompensated factor pressure, ψ/ψ_0 , and the available AE. Kinetic constants are denoted by $k_{...}$. The phase portrait of this model is completely determined by the position of the two lines, where the time derivative of r_0 or r vanishes (nullclines) (see Fig. 28 where for certainty we assume that $k_{pr} > k$, which guarantees the concavity of the r_0 nullcline; the r nullcline is always convex).

The vertical line $W = 0$ (that is, $f - r = \psi_0$) is the border of death. If $f < \psi_0$ then the r -nullcline does not intersect the border of death, there exists a globally stable equilibrium and the factor pressure cannot lead to death (Fig. 28a). If $f > \psi_0$ then a dangerous (absorbing) fragment appears on the border of death (Fig. 28b,c, in red). Up to a certain threshold of the factor pressure, the nullclines of r and r_0 intersect, and there are two stationary states: a (locally) stable node and a saddle point, the separatrix of which is the boundary between the basins of attraction of a stable node and of the dangerous (absorbing) part of the death border (Fig. 28b). Further increase of the factor pressure f leads to the second critical effect: intersection between nullclines vanishes and the absorbing part of the border of death becomes globally attractive (Fig. 28c).

Many physiologists and clinicians note the fact that when the adaptation load increases and approaches a certain threshold, a jump in adaptive abilities occurs. Selye and his successors found that just the idea of AE supply cannot explain many such effects. Selye introduced the notion of “deep AE” (deep here means not another kind of AE, but rather its deeper storage) and used for the routinely accessible AE the term “superficial AE”. The fourth axiom of AE reads (accordingly to [143]):

- AE is active at two levels of awareness: a primary level at which creating the response occurs at a high awareness level, with high usage of finite supply of adaptation energy; and a secondary level at which the response creation is being processing at a sub-awareness level, with a lower energy expenditure.

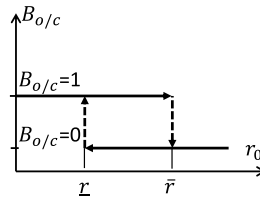


Fig. 29. Resource – reserve hysteresis. Hysteresis of reserve supply: if $B_{o/c} = 0$ then reserve is closed and if $B_{o/c} = 1$ then reserve is open. When r_0 decreases and approaches $\underline{r}$ then the supply or reserve opens (if it was closed). When $r_0 < \bar{r}$ increases and approaches $\bar{r}$ then the supply of reserve closes (if it was open).

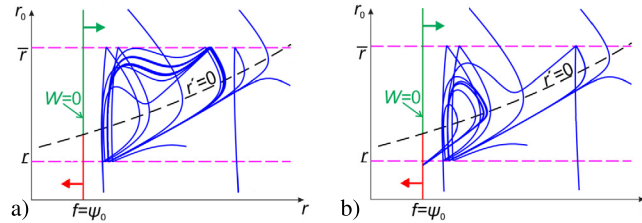


Fig. 30. Oscillations near the border of death for 3D subsystem in projection onto the (r, r_0) plane (the third coordinate, r_{rv} , is hidden). (a) Stable oscillations; (b) Oscillating death. Several trajectories for different initial states are plotted together for each case (a), (b). Obtained by numerical integration of the model. For detail and more examples see [72].

This deep AE level is represented in the model diagram (Fig. 27) by the AE reserve (left). There are two storages of AE: resource (which is always available if it is not empty) and reserve (which becomes available when the resource becomes too low). The Boolean variable $B_{o/c}$ describes the state of the ‘tap’ between the reserve and the main storage: if $B_{o/c} = 0$ then the tap is closed and if $B_{o/c} = 1$ then the tap is open. There are two switch lines on the phase plane (r, r_0) (Fig. 29): $r_0 = \underline{r}$ (the lower switch line that serves to opening the reserve storage) and $r_0 = \bar{r}$ (the upper switch line that serves to closing the reserve storage).

When the available resource r_0 decreases and approaches $\underline{r}$ from above then the supply or reserve opens (if it was closed). When the available resource $r_0 > \bar{r}$ increases and approaches $\bar{r}$ from below then the supply of reserve closes (if it was open). For $r_0 < \underline{r}$ the reserve is always open, $B_{o/c} = 1$ and for $r_0 > \bar{r}$ the reserve is always closed, $B_{o/c} = 0$.

A system of three blocks of the scheme shown in Fig. 27, from left to right: reserve, resource and allocated resource, is a 3D dynamic system with rich behavior. It seems that Selye and other physiologists underestimated the dynamic consequences of existence of additional reserve AE storage with on and off thresholds. These thresholds lead to oscillations. For example, there are the stable closed orbits (Fig. 30a). If we decrease the reserve recovering constant k_{pr1} then the closed orbits lose stability and the oscillating death appears (Fig. 30b). More examples are presented in [72].

Fig. 30b demonstrates an important effect: the trajectories spend a long time near the places where cycles appear for different values of constants (see Fig. 30a) and go to the attractor (here it is death) after this delay. The delayed relaxation is a manifestation of the so-called ‘critical retardation’: near a bifurcation with the appearance of new ω -limit points, the trajectories spend a long time close to these points [62].

The phenomena of the oscillating remission and oscillating death have been observed in clinic for a long time and now attract attention in mathematical medicine and biology. For example, [181] demonstrated recently, on a more detailed model of adaptation in the immune system, that cycles of relapse and remission, typical for many autoimmune diseases, arise naturally from the dynamical behavior of the system. The notion of ‘oscillating remission’ is used also in psychiatry [74].

For extracting oscillating death and recovering from statistical data, synchronization is useful. Without phase synchronization, averaging smears the periodic component. Synchronization helps to extract the periodicity. There exist some medical data sets with obvious synchronization. For example, in trauma, there exists a selected moment – the trauma event. Of course, in severe injuries, mortality is high at the beginning, and this large peak with rapid decay shows no fluctuations. However, for low-severity injuries (for example, for NISS severity scores 1–8 or for NISS severity scores 9, where NISS = New Injury Severity Score), the daily mortality rate – the estimated probability of

a patient's death per day t , provided that he survived for a day $t - 1$, is definitely not a monotonous function of time and includes a pair of periods of decaying fluctuations with a period of 12–16 days. This was demonstrated by [114] on data from TARN - the Trauma Audit & Research Network that has the largest trauma database in Europe. Another example is the oscillating mortality after cancer surgery (high severity cohort) [103]. Here the sync event is an operation.

The models based on Selye's idea of deep AE demonstrate the oscillating remission and oscillating death. These effects do not need additional exogenous reasons and are the core effects of adaptation process. In further research [55] additional levels of adaptation processes were introduced at a lower intensity of stressors, and a “periodic table” of adaptation reactions was developed. According to them, the ‘chain of AE supply’ in Fig. 27 should be extended to the left to many storages and taps with thresholds.

Despite the interesting dynamical effects, the AE resource and reserve model (three left blocks in Fig. 27) has an important drawback. They do not describe training. Training means, in particular, that the possible factor load that organism can resist may be increased after some experience.

For example, in Selye's experiment, rats pretreated with a certain agent will resist such doses of this agent in the future. At the same time, their resistance to toxic doses of that one decreases [148]. That is, the resistance to noxious agents can be trained and the training effect lasts for a rather long time. At the same time, this training decreases the ability to resist other agents. In a recent review, the applicability of the concept of AE to the training process was questioned [169]: “It seems unlikely that adaptation energy can explain adaptation and acute responses to exercise. Instead, the concept of adaptation energy appears to share similarities with some theories of aging.” We agree that the previously developed models do not describe training, but the righthmost block of the model Fig. 27 aims to solve this problem. AE is spending not only for neutralization of the harmful factor but to extending the zone of surviving measured by the logarithm of its volume, ψ_0 . The logarithm of volume in the factor space where life is possible can be considered as entropy and training is an extension of this volume. The new terms are simple and it is worth to look on them separately. (For the complete system, see Fig. 27 and Appendix C.):

$$\begin{aligned} \frac{dr}{dt} &= \dots - k_{tr} r \frac{\psi}{\psi_0} W_0 \left(\frac{\psi}{\psi_0} \right); \\ \frac{d \ln \psi_0}{dt} &= -k_{detr} + \alpha k_{tr} r \frac{\psi}{\psi_0} W_0 \left(\frac{\psi}{\psi_0} \right), \end{aligned} \quad (10)$$

where α is the coefficient of training efficiency ($0 < \alpha < 1$), $W_0(x) = 1 - x$ for $0 \leq x \leq 1$.

The main assumptions behind this entropic approach are:

- Training is a medium-term adaptation process with transformation of the AE into entropic form, $\ln \psi_0$.
- Following Selye, this transformation is irreversible and the transformed energy cannot be spent to other purposes.
- Life length of this entropic form is determined by the detraining constant k_{detr} .

Mathematical modeling of athletic training and performance attracted much attention in physiology of sport and exercises. The models are considered as effective conceptual and practical tools for evidence-based exercise (see the review [34]). Performance models are used as the basis for training design and training planning. We expect some analogy between these models and the learning process models that can be extracted from the model Fig. 27 and (10), in particular. One of the key concepts in training models is ‘Critical Power’ (CP), that is, the maximum speed of work that can theoretically be performed for an infinite time and corresponds to the maximum sustained aerobic power. During exercise at power above CP, there is a clear and progressive loss of metabolic homeostasis. This CP can be related to our ψ_0 . It can be worth to mention that for humans characteristic detraining time is evaluated as $1/k_{detr} \approx 8 - 10$ weeks in some athletic training exercises [34].

A natural question arises: if the resource allocation r is fixed, what is the relationship between the training load ψ and the dynamics ψ_0 , which measures the ability to survive under uncompensated pressure of the harmful factor? The answer is given by the nullcline of the equation for ψ_0 (10): The positive effect of the training process (the growth of ψ_0) holds for

$$\frac{1}{2} - \sqrt{\frac{1}{4} - \frac{k_{detr}}{\alpha k_{tr} r}} \leq \frac{\psi}{\psi_0} \leq \frac{1}{2} + \sqrt{\frac{1}{4} - \frac{k_{detr}}{\alpha k_{tr} r}} \quad (11)$$

if the available resource $r > 4k_{detr}/\alpha k_{tr}$. If this inequality does not hold then training is impossible and $d\psi_0/dt \leq 0$ for all values of ψ . For a positive training effect (growth ψ_0), the uncompensated factor pressure ψ must be in the interval (11) with the midpoint at $\psi_0/2$. The width of this interval depends on the allocated resource r . If this resource decreases, the interval width decreases and becomes zero at some final value of r . After that, the training may just slow down the decline of ψ_0 . The optimal training load ψ remains half of the maximum load ψ_0 . This example demonstrates the applicability of the model presented in Fig. 27 to qualitative modeling of training. To obtain accurate quantitative results, it is necessary to work out in more detail the interpretation of the model in directly observed quantities and the estimation of constants.

5. Conclusion and outlook

In many works of various groups of researchers, it has been shown that the dynamics of correlations and variances in the ensembles of systems that adapt to the pressure of harmful factors has universal properties: under stress, both correlations and variance increase. This behavior lasts till some critical intensity of stressors. After that, the correlations may decrease. This effect is supported by many observations in ecological physiology, medicine, economics and finance and may serve for early diagnosis of crises. Analyzing the correlation graph between attributes provides a tool for studying adaptation, stress and critical transitions.

These effects appear for sets of relevant attributes. These attributes can be different for different factors. Selecting the appropriate attributes is an important part of the analysis. There are several approaches to this choice. It can be done on the basis of preliminary hypotheses about the adaptation mechanisms or attributes with higher relative change of their values can be selected. The structure of the correlation graph can also be used with heuristics: internal connections in a set of relevant attributes under stress are (on average) stronger than external connections.

We presented several examples of the opposite behavior. The reasons for such deviations may be different, from the unsatisfactory choice of the relevant attributes to the special interactions in the system of factors, like synergy. In any case, these examples warn us that effect validation should be performed for any type of system, harmful factors, and attribute selection before practical application.

Several attempts have been made to develop a theory for this effect. We focused here on the classical physiological idea about adaptation resource (AE) expressed by Selye. Combined with the idea of limitation, this approach explains the observations and predicts other effects for different organization of systems of factors (for synergistic systems of factors). It also allows to describe the effect of training in a natural way. Other approaches, such as the idea of bifurcations, critical retardations and critical fluctuations, also explain some effects but still fail to explain phenomena that were described by Selye as spending and exhausting of adaptation resource.

The idea of AE and its optimal distribution to neutralize harmful factors has analogy with thermodynamics. These general thermodynamic models are supplemented by kinetic models of the production and expenditure of AE in individual adaptation.

Selye's 'AE' is an abstract adaptation resource, the universal currency for adaptation. Two types (at least) of the adaptation resource supply are needed for this modeling: from 'checking account' of the available resource (the superficial energy) and from 'saving account' of the reserve (the deep energy). Existence of these two types determine rich family of dynamical regimes including limiting cycles and oscillating death.

Analysis of training models leads to introduction of *entropic* form of adaptation energy. This energy is irreversibly pumped into the entropic reservoir, extends the volume of comfort zone (this is the training effect) and cannot be reassigned. Models with AE and adaptation entropy capture main phenomenological effects and can be used in the top-down modeling of physiological adaptation. This new quantity has a transparent geometric interpretation. Consider the area in the space of factor values, where the organism can survive without pumping AE. Formally, this 'area of survival' can be defined by positivity of fitness W , as an area of ψ where $W(\psi) > 0$. Let the volume of this area be V . Then the AE pumped in the entropic reservoir is $\ln V$.

Our goal was to understand the classical experiments and to build the basic top-down models of physiological adaptation. Now, after 33 years of the first publication [68] and efforts of many dozens of researchers, the main problems become clear.

Let us first mention applications. Application practice will change the systems of models and create new questions. Three groups of applications should be noted:

- Development of early warning indicators. The potential of using the correlation graph and its characteristics is clearly demonstrated for early warning of stress and crisis for various systems, from individual health to the health system, social tensions and financial crashes (see Sec. 2).
- Application of the models of individual adaptation (Sec. 4) to training and sport physiology. These models should be incorporated into the existing model system, possibly with modifications, and used for training design and planning.
- Detailed analysis of the oscillating recovery and oscillating death. These phenomena are recently predicted by the Selye–Goldstone kinetic AE models (Sec. 4) [72] and give a key to dynamic of individual adaptation and new branch of predictive personalized medicine. The kinetic models provide a tool for description, prediction, and work with these phenomena.
- Application of the dynamic and thermodynamic models of adaptation to the analysis of aging and development of new methods of optimal aging control. Healthy and successful aging, is not connected with excessive capacity to cope with stressors but rather those with capability to afford intermediate response to stress, i.e. neither too weak nor too strong, in relation to the environment [51]. In terms of AE distribution, this moderation can be seen as optimal use of AE for the purpose of longevity.
- Development of specific models of various systems, such as cell cultures and isolated organs, that allow detailed experimental testing. Using these models in planning experiments. A first example is a simple model of biological neural networks was developed that explains the generation of bursts [106] in the networks of living culture [166]. Each neuron is characterized by a dynamic exogenous energy variable and neuronal activation probability depends on the energy and on the external activating signals from the network.

Secondly, the fundamental problem of interpretation and physical meaning of AE has been around for a long time. Selye introduced this concept as ‘adaptability’, and the analogy with thermodynamics highlighted his idea, but was not confirmed by physical experiments or solid theory [147,148]. This elusiveness of the AE raised many questions and doubts. In our research, we considered this quantity as a coordinate on the ‘dominant path’ of adaptation processes and avoided direct physical interpretation [72].

Surprisingly, the physical idea of AE (or rather, free adaptation energy) has recently been revived in the physical theory of adaptive systems [2,94]. The relationship between Selye’s AE and the expenditure of physical free energy on adaptation should be clarified. Now we can hypothesize that this physical quantity can be considered as the lower bound of the adaptation energy, just as the physical efficiency coefficient is the upper bound of the efficiency coefficients of real devices.

The third group of problems is related to the dynamics of correlations and variances ‘on another side of crisis’, Fig. 3. The cases near fatal outcome are explored in much less detail than the initial stages of stresses, diseases and crises of adaptable systems. More attention is needed for this area because understanding of death is necessary for deep understanding of life.

Detailed parametric analysis of phase portraits for dynamics of individual adaptation with training Fig. 27 and for the simpler resource–reserve dynamics will answer many practically important question about general adaptation syndrome that remain open for decades [57,58]. On the other hand, this analysis can bring new ideas and benchmarks to dynamics of adaptive systems [165].

Development of optimal control methods for adaptation dynamics will create a new discipline, “adaptation engineering”, a special branch of applied physiology and personalized preventive health care.

Analysis and classification of systems of factors requires a lot of work. Special attention should be paid to such *universal physiological factors* as inflammation, which can be considered as a universal mechanism for accelerating aging (‘inflammaging’ [50]). The model “universal resource-system of universal factors” that describes the dynamics of the spending of AE for protection from several universal harmful factors can be useful in various areas of physiology. A simple quasistatic (‘thermodynamic’) theory of factor’s interactions is presented in Section 3.2. For the future applications to monitoring and control of individual adaptation, it is necessary to develop dynamic models of the interaction of several factors and AE. The problem here is that there are too many possibilities for a formal description of this interaction and additional empirical clues are needed.

Finally, a systematic search for counterexamples to the main effect is required. The analysis of these contradictory cases will give us new ideas and will stimulate the development of theory and the discrimination between different theoretical ideas. It is now possible to indicate several potential reasons for the violation of the main effect (see Fig. 2):

- Inappropriate choice of the relevant variables;
- Approaching the ‘point of no return’ on another side of crisis (see Fig. 3);
- Synergy between several harmful factors (Sec. 3.5).

The analysis of counterexamples will help in solving a really complex and difficult conceptual problem: choosing relevant attributes. There are many plausible algorithms and heuristics for identifying these variables, from simple common sense, intuition, and professional experience to some formalized algorithms with several tunable parameters. In real multidimensional problems, the variety of solutions can be enormous.

This problem should not be considered as purely technical. It has an intrinsic connection with the deep problem of representing a system (an organism or other adaptable system) as a construction of relatively independent functional subsystems. Factor analysis and principal component analysis are used to identify such subsystems. Independent component analysis (ICA) was specially invented for separation of a complex signal into signals from independent subsystems [82] (For the technical aspects of ICA applications such as determining the optimal number of components and improving reproducibility of the results see [156]). The theory of functional systems in biology was developed before the big data revolution [3,4,160]. The configuration of subsystems is dynamic and changes over time and under the pressure of various factors. Now is the time to revive functional systems theory as a data-driven discipline closely related to machine learning [137].

The selection of relevant features can have a profound sense of revealing relevant subsystems. Representation of the organism as a ‘universe’ of flexibly joining and splitting functional subsystems is very attractive. It could be suitable to cite here the old wisdom: “The land under heaven, after a long period of division, tends to unite; after a long period of union, tends to divide.” (The “Romance of the Three Kingdoms”, a historical epic fiction written in the fourteenth century.)

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The work was supported by the University of Leicester and the Ministry of Science and Higher Education of the Russian Federation (Project No. 14.Y26.31.0022).

Appendix A. Thermodynamic cost of homeostasis

Selye experiments added a seminal idea to the homeostasis control loops: adaptation has a cost. Surprisingly, a simple thermodynamic estimate of the cost of homeostasis in the terms of free energy is possible.

Consider a non-equilibrium physicochemical system. Let its internal state be described by a vector of extensive variables N . If the system is closed and free energy is not added, then it goes to equilibrium. For living systems, this necessarily means death. For dynamics of the system without external fluxes we can write a general dynamic equation:

$$\frac{dN}{dt} = W(N). \quad (\text{A.1})$$

This equation must satisfy the first and second laws of thermodynamics: the conservation laws must be satisfied, and the production of entropy must be positive. For isothermal conditions the second law means a decrease in free energy: the convex function $G(N)$ is defined and

$$\frac{dG}{dt} \leq 0, \quad (\text{A.2})$$

the time derivative of G by virtue of the system (A.1) is non-positive. Let us denote this decrease rate of G in the closed system by $-\sigma(N)$:

$$\sigma(N) = -(\nabla G|_N, W(N)) \geq 0.$$

Suppose that we aim to stabilize the non-equilibrium state N^{st} by external fluxes. For this purpose, N^{st} should be a steady state of the system (A.1) with added external fluxes:

$$\frac{dN}{dt} = W(N) + v(N_{\text{in}} - N), \quad (\text{A.3})$$

where v is an intensity of flux, and N_{in} is a vector of input composition. If N^{st} is a steady state of this system then the time derivative of G by virtue of the system (A.3) must vanish. This means that

$$\sigma(N^{\text{st}}) = v(\nabla G|_{N^{\text{st}}}, N_{\text{in}} - N^{\text{st}}). \quad (\text{A.4})$$

According to Jensen's inequality for convex function G ,

$$(\nabla G|_{N^{\text{st}}}, N_{\text{in}} - N^{\text{st}}) \leq G(N_{\text{in}}) - G(N^{\text{st}}). \quad (\text{A.5})$$

Combining (A.4) and (A.5) we obtain the inequality

$$v(G(N_{\text{in}}) - G(N^{\text{st}})) \geq \sigma. \quad (\text{A.6})$$

The left part of (A.6) is the inflow of free energy into the system, and the right part is the dissipation rate for a closed system in the same state N^{st} (this state is, of course, the transition state of the closed system).

That is, the inflow of free energy required to stabilize the N^{st} state in an open system must be greater than or equal to the rate of dissipation of free energy in a closed system at the same state N^{st} .

It should be stressed that the inequality (A.6) is necessary but not sufficient condition for stabilization. Various versions of this inequality are used in many works about thermodynamics of homeostasis.

Appendix B. Advection-diffusion model of adaptation dynamics near a border of existence

We used the idea of adaptation resource to demonstrate how can the models of adaptation explain the observable dynamics of correlation and variance under environmental loads. Another theoretical approach to explain these effects was proposed in [135] and developed further with applications in [136,150]. The idea is quite simple and attractive. Let us represent the population of systems by points in a bounded set $U \subset \mathbb{R}^q$ with smooth boundary. Coordinates correspond to the internal characteristics of the systems. The whole population is described by a continuous density function $u(x)$ in U .

The hypothetical behavior of the individual system can be described as a random walk in U , and the density dynamics satisfies the correspondent Fokker–Planck equation

$$\partial_t u = -(\nabla, bu) + a\Delta u, \quad (\text{B.1})$$

where $a > 0$ is a constant diffusion coefficient and b is a drift vector. This is a simple advection-diffusion equation with the constant advection speed b . After analysis of various optimality conditions, the asymptotic behavior of solution of Eq. (B.1) was studied for $b = fb_0$, $b_0 = \text{const}$ ($\|b_0\| = 1$), $f \rightarrow \infty$, and for no-flux boundary conditions $(bu - a\nabla u, \nu)_{\partial U} = 0$ (where ν is the outer normal to ∂U). The parameter f models the pressure of environmental factors. For large f , the drift presses the distribution to the boundary of U , dimension decreases, and correlations increase. This picture is schematically represented in Fig. 26.

The asymptotics for large f is determined by a tangent paraboloid to ∂U at the point x_b , where the linear function (b, x) achieves the maximal value on ∂U (generically, such a point is unique). At this point $\nu = b_0$. The asymptotic of the solution coincides in the main term with exact analytical solution to the Fokker–Planck equation (B.1) with constant coefficients in a paraboloid.

Let us select a rectangular coordinate system with origin at the point x_b , such that in this system $b = (0, \dots, 0, -A)$ and U is given by the inequality

$$x_q \geq \sum_{i=1}^{q-1} \lambda_i x_i^2 + o\left(\sum_{i=1}^{q-1} x_i^2\right).$$

In this coordinate system, the solution to the problem is

$$u(x) = v \exp(-fx_q/a)$$

with constant v . For large A , the variance of x_q , $\sigma_q^2 \sim a^2/f^2$, the marginal distribution of $x_1, \dots, x_{q-1}$ is almost Gaussian with $\sigma_i^2 = a/f\lambda_i$, and these variables are almost independent. We can see that the concentration ellipsoid for $x_1, \dots, x_{q-1}$ does not change its form with increase of A : the ratio $\sigma_i^2/\sigma_j^2 = \lambda_j/\lambda_i$ is independent of f . On the other hand, contraction in direction x_q goes faster, and $\sigma_q^2/\sigma_i^2 \rightarrow 0$ as $a\lambda_i/f$ ($i = 1, \dots, q-1$).

To check which mechanism is closer to reality we have to resolve the “disk-or-bullet” alternative: How does the correlation increase: (i) the largest eigenvalues of the correlation matrix increase and the ellipsoid of concentration becomes stretched-out in small amount of dimensions (presumably, in one, that is the bullet form) or (ii) the smallest eigenvalue goes fast to zero, and the ellipsoid of concentration becomes flattened out (the disk or the pancake form). The existing data ensure us that this alternative should be resolved in favor of bullet: under the pressure of environment factors, the small amount of largest eigenvalues increases. It is necessary to mention that the variance decreases in the Fokker–Planck model, that also contradicts the majority of existing observation.

Appendix C. AE kinetic equations

Here, we present the system of equations that correspond to the scheme Fig. 27. The main variables are three amounts of AE: r – AE allocated for factor neutralization; r_0 – AE available at low level of alarm (‘superficial’), r_{rv} – AE accumulated in reserve, and $\beta \ln \psi_0$ – entropic form of AE, proportional to logarithm of the area of surviving in the space of factor values. The system is presented below:

$$\begin{aligned} \frac{dr}{dt} &= -k_d r + k r_0 \frac{\psi}{\psi_0} - k_{tr} r \frac{\psi}{\psi_0} W_0 \left(\frac{\psi}{\psi_0} \right); \\ \frac{dr_0}{dt} &= -k_{d0} r_0 - k r_0 \frac{\psi}{\psi_0} + k_{rv} B_{o/c} r_{rv} (R_0 - r_0) + k_{pr} (R_0 - r_0) W_0 \left(\frac{\psi}{\psi_0} \right); \\ \frac{dr_{rv}}{dt} &= -k_{d1} r_{rv} - k_{rv} B_{o/c} r_{rv} (R_0 - r_0) + k_{pr1} (R_{rv} - r_{rv}) W_0 \left(\frac{\psi}{\psi_0} \right); \\ \frac{1}{\psi_0} \frac{d\psi_0}{dt} &= -k_{detr} + \alpha k_{tr} r \frac{\psi}{\psi_0} W_0 \left(\frac{\psi}{\psi_0} \right), \end{aligned} \quad (C.1)$$

where $W_0(x) = 1 - x$ for $0 \leq x \leq 1$ (this is the scale-invariant form of W), α is the coefficient of efficiency of this training ($0 < \alpha < 1$), $k_{\dots}$ are the rate constants. The Boolean variable $B_{o/c}$ describes the state of the tap between the reserve and the resource storages (Fig. 29).

System (C.1) formalizes the main general ideas of Selye, Goldstone and other physiologists and describes phenomena related to dynamics of adaptation to a harmful factor. The detailed parametric analysis of this system and evaluation of its parameters for real life examples is now a priority task. Goldstone’s ‘GP challenge’ to the theory of general adaptation syndrome [57,58] should be met. To meet them completely, we must also study the dynamic interaction of adaptation to several factors. However, to create such dynamic models, it is necessary to analyze, identify, and validate the dynamic model of adaptation to a single factor (C.1) in more detail.

References

- [1] Adjero D, Cao D, Piccirilli M, Ross A. Predictability and correlation in human metrology. In: Proc. of 2010 IEEE international workshop on information forensics and security (WIFS2010). IEEE Press; 2010. p. 1–6.
- [2] Allahverdyan AE, Babajanyan SG, Martirosyan NH, Melkikh AV. Adaptive heat engine. Phys Rev Lett 2016;117(3):030601. <https://doi.org/10.1103/PhysRevLett.117.030601>.
- [3] Anokhin PK. Systemogenesis as a general regulator of brain development. Prog Brain Res 1964;9:54–86. [https://doi.org/10.1016/S0079-6123\(08\)63131-3](https://doi.org/10.1016/S0079-6123(08)63131-3).
- [4] Anokhin PK. Nodular mechanism of functional systems as a self-regulating apparatus. Prog Brain Res 1968;22:230–51. [https://doi.org/10.1016/S0079-6123\(08\)63509-8](https://doi.org/10.1016/S0079-6123(08)63509-8).
- [5] Aoyama H, Aruka Y, Yoshikawa H, editors. Complexity, heterogeneity, and the methods of statistical physics in economics. Evolutionary economics and social complexity science, vol. 22. Singapore: Springer; 2020.

- [6] Armbruster WS, Di Stilio VS, Tuxill JD, Flores TC, Runk JL. Covariance and decoupling of floral and vegetative traits in nine Neotropical plants: a re-evaluation of Berg's correlation-pleiades concept. *Am J Bot* 1999;86(1):39–55. <https://doi.org/10.2307/2656953>.
- [7] Arnold SJ. Morphology, performance and fitness. *Am Zool* 1983;23(2):347–61. <https://doi.org/10.1093/icb/23.2.347>.
- [8] Aumont O, Maier-Reimer E, Blain S, Monfray P. An ecosystem model of the global ocean including Fe, Si, P colimitations. *Glob Biogeochem Cycles* 2003;17(2):1060. <https://doi.org/10.1029/2001GB001745>.
- [9] Barik S, Das S, Vikalo H. QSDpR: viral quasispecies reconstruction via correlation clustering. *Genomics* 2018;110(6):375–81. <https://doi.org/10.1016/j.ygeno.2017.12.007>.
- [10] Berg RL. The ecological significance of correlation pleiades. *Evolution* 1960;1:171–80. Available from: <http://www.jstor.org/stable/2405824>.
- [11] Bernardini C, Censi F, Lattanzi W, Barba M, Calcagnini G, Giuliani A, et al. Mitochondrial network genes in the skeletal muscle of amyotrophic lateral sclerosis patients. *PLoS ONE* 2013;8(2):e57739. <https://doi.org/10.1371/journal.pone.0057739>.
- [12] Billman GE. Homeostasis: the underappreciated and far too often ignored central organizing principle of physiology. *Front Physiol* 2020;11:200. <https://doi.org/10.3389/fphys.2020.00200>.
- [13] Bloom AJ, Chapin III FS, Mooney HA. Resource limitation in plants – an economic analogy. *Annu Rev Ecol Syst* 1985;16:363–92. <https://doi.org/10.1146/annurev.es.16.110185.002051>.
- [14] Bomze IM. Regularity vs. degeneracy in dynamics, games, and optimization: a unified approach to different aspects. *SIAM Rev* 2002;44(3):394–414. <https://doi.org/10.1137/S00361445003756>.
- [15] Brenner S. Biological computation. In: Bock GR, Goode JA, editors. *The limits of reductionism in biology, novartis foundation symposium*, vol. 213. Chichester: Wiley; 1998. p. 106–16.
- [16] Brillinger DR. Remarks concerning graphical models for time series and point processes. *Rev Econom/SBE* 1996;16:1–23. <https://doi.org/10.12660/bre.v16n11996.2878>.
- [17] Breznitz, editor. *The denial of stress*. New York: International Universities Press, Inc.; 1983.
- [18] Bulygin GV, Mansurov AS, Mansurova TP, Smirnova EV. Dynamics of parameters of human metabolic system during the short-term adaptation. Preprint 180B. Institute of Biophysics, Russian Academy of Sciences; 1992.
- [19] Bulygin GV, Mansurov AS, Mansurova TP, Mashanov AA, Smirnova EV. Impact of health on the ecological stress dynamics. Preprint 185B. Krasnoyarsk: Institute of Biophysics, Russian Academy of Sciences; 1992.
- [20] Butte AJ, Kohane IS. Unsupervised knowledge discovery in medical databases using relevance networks. In: *Proc. AMIA symp. American Medical Informatics Association*; 1999. p. 711–5. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/pmc2232846/>.
- [21] Butte AJ, Kohane IS. Mutual information relevance networks: functional genomic clustering using pairwise entropy measurements. In: *Proc. of the Pacific symposium biocomputing 2000*; 1999. p. 418–29.
- [22] Cangelosi R, Goriely A. Component retention in principal component analysis with application to cDNA microarray data. *Biol Direct* 2007;2:2. <https://doi.org/10.1186/1745-6150-2-2>.
- [23] Cannon WB. *The wisdom of the body*. NY: W.W. Norton; 1932.
- [24] Carpenter RHS. Homeostasis: a plea for a unified approach. *Adv Physiol Educ* 2004;28(4):180–7. <https://doi.org/10.1152/advan.00012.2004>.
- [25] Censi F, Calcagnini G, Bartolini P, Giuliani A. A systems biology strategy on differential gene expression data discloses some biological features of atrial fibrillation. *PLoS ONE* 2010;5(10):e13668. <https://doi.org/10.1371/journal.pone.0013668>.
- [26] Censi F, Giuliani A, Bartolini P, Calcagnini G. A multiscale graph theoretical approach to gene regulation networks: a case study in atrial fibrillation. *IEEE Trans Biomed Eng* 2011;58(10):2943–6. <https://doi.org/10.1109/TBME.2011.2150747>.
- [27] Censi F, Calcagnini G, Mattei E, Giuliani A. System biology approach: gene network analysis for muscular dystrophy. In: Bernardini C, editor. *Duchenne muscular dystrophy. Methods in molecular biology*, vol. 1687. New York, NY: Humana Press; 2018. p. 75–89.
- [28] Chakrabarti AS. Network theory in macroeconomics and finance. In: *Econophysics and sociophysics: recent progress and future directions*. Cham: Springer; 2017. p. 71–83.
- [29] Chapin III FS, Schulze E, Mooney HA. The ecology and economics of storage in plants. *Annu Rev Ecol Syst* 1990;21:423–47. <https://doi.org/10.1146/annurev.es.21.110190.002231>.
- [30] Chappell L, Russell AJ, Voet T. Single-cell (multi) omics technologies. *Annu Rev Genomics Hum Genet* 2018;19:15–41. <https://doi.org/10.1146/annurev-genom-091416-035324>.
- [31] Chen H, Albergante L, Hsu JY, Lareau CA, Bosco GL, Guan J, et al. Single-cell trajectories reconstruction, exploration and mapping of omics data with STREAM. *Nat Commun* 2019;10(1):1903. <https://doi.org/10.1038/s41467-019-09670-4>.
- [32] Chen L, Liu R, Liu ZP, Li M, Aihara K. Detecting early-warning signals for sudden deterioration of complex diseases by dynamical network biomarkers. *Sci Rep* 2012;2(342). <https://doi.org/10.1038/srep00342>.
- [33] Chiang TC, Jeon BN, Li H. Dynamic correlation analysis of financial contagion: evidence from Asian markets. *J Int Money Financ* 2007;26(7):1206–28. <https://doi.org/10.1016/j.jimonfin.2007.06.005>.
- [34] Clarke DC, Skiba PF. Rationale and resources for teaching the mathematical modeling of athletic training and performance. *Adv Physiol Educ* 2013;37(2):134–52. <https://doi.org/10.1152/advan.00078.2011>.
- [35] Cramer AOJ, van Borkulo CD, Giltay EJ, van der Maas HLJ, Kendler KS, Scheffer M, et al. Major depression as a complex dynamic system. *PLoS ONE* 2016;11(12):e0167490. <https://doi.org/10.1371/journal.pone.0167490>.
- [36] Danger M, Daufresne T, Lucas F, Pissard S, Lacroix M. Does Liebig's law of the minimum scale up from species to communities?. *Oikos* 2008;117(11):1741–51. <https://doi.org/10.1111/j.1600-0706.2008.16793.x>.
- [37] de Felice G, Giuliani A, Halfon S, Andreassi S, Paoloni G, Orsucci FF. The misleading Dodo Bird verdict. How much of the outcome variance is explained by common and specific factors?. *New Ideas Psychol* 2019;54:50–5. <https://doi.org/10.1016/j.newideapsych.2019.01.006>.
- [38] de Felice G, Giuliani A, Gelo OCG, Mergenthaler E, De Smet MM, Meganck R, et al. Differentiates poor- and good-outcome psychotherapy? A statistical-mechanics-inspired approach to psychotherapy research, part two: network analyses. *Front Psychol* 2020;11:788. <https://doi.org/10.3389/fpsyg.2020.00788>.

- [39] de Felice G, Orsucci FF, Scozzari A, Gelo O, Serafini G, Andreassi S, et al. What differentiates poor and good outcome psychotherapy? A statistical-mechanics-inspired approach to psychotherapy research. *Systems* 2019;7:22. <https://doi.org/10.3390/systems7020022>.
- [40] Darchambeau F, Faerøvig PJ, Hessen DO. How *Daphnia* copes with excess carbon in its food. *Oecologia* 2003;136:336–46. <https://doi.org/10.1007/s00442-003-1283-7>.
- [41] Ding KF, Finlay D, Yin H, Hendricks WP, Sereduk C, Kiefer J, et al. Network rewiring in cancer: applications to melanoma cell lines and the cancer genome atlas patients. *Front Genet* 2018;9:228. <https://doi.org/10.3389/fgene.2018.00228>.
- [42] Dobzhansky T. Evolution in the tropics. *Am Sci* 1950;38(2):209–21. Available from: <https://www.jstor.org/stable/27826306>.
- [43] Drożdż S, Grümm F, Górski AZ, Ruf F, Speth J. Dynamics of competition between collectivity and noise in the stock market. *Physica A* 2000;287:440–9. [https://doi.org/10.1016/S0378-4371\(00\)00383-6](https://doi.org/10.1016/S0378-4371(00)00383-6).
- [44] Dupré JE. *The latest on the best: essays on evolution and optimality*. The MIT Press; 1987.
- [45] Dykman MI, Krivogla MA. Fluctuations in nonlinear systems near bifurcations corresponding to the appearance of new stable states. *Physica A* 1980;104(3):480–94. [https://doi.org/10.1016/0378-4371\(80\)90010-2](https://doi.org/10.1016/0378-4371(80)90010-2).
- [46] Egli T, Zinn M. The concept of multiple-nutrient-limited growth of microorganisms and its application in biotechnological processes. *Biotechnol Adv* 2003;22(1–2):35–43. <https://doi.org/10.1016/j.biotechadv.2003.08.006>.
- [47] Ehrlich E, Becks L, Gaedke U. Trait–fitness relationships determine how trade-off shapes affect species coexistence. *Ecology* 2017;98(12):3188–98. <https://doi.org/10.1002/ecy.2047>.
- [48] Fehrman E, Egan V, Gorban AN, Levesley J, Mirkes EM, Muhammad AK. Personality traits and drug consumption. A story told by data. Cham: Springer; 2019. Available from: <https://arxiv.org/abs/2001.06520>.
- [49] Fisher RA. *The genetical theory of natural selection*. Oxford: Oxford University Press; 1930. Available from: <https://archive.org/stream/geneticaltheory031631mbp/page/n11/mode/2up>.
- [50] Franceschi C, Bonafè M, Valensin S, Olivieri F, De Luca M, Ottaviani E, et al. An evolutionary perspective on immunosenescence. *Ann NY Acad Sci* 2000;908(1):244–54. <https://doi.org/10.1111/j.1749-6632.2000.tb06651.x>.
- [51] Franceschi C, Valensin S, Bonafè M, Paolisso G, Yashin AI, Monti D, et al. The network and the remodeling theories of aging: historical background and new perspectives. *Exp Gerontol* 2000;35(6–7):879–96. [https://doi.org/10.1016/S0531-5565\(00\)00172-8](https://doi.org/10.1016/S0531-5565(00)00172-8).
- [52] Fried R, Didelez V, Lanius V. Partial correlation graphs and dynamic latent variables for physiological time series. In: Baier Daniel, et al., editors. *Innovations in classification, data science, and information systems, proc. of the 27th annual conference of the Gesellschaft für Klassifikation e. V.* Berlin: Springer; 2005. p. 259–66.
- [53] Fukunaga K, Olsen DR. An algorithm for finding intrinsic dimensionality of data. *IEEE Trans Comput* 1971;C-20:176–83. <https://doi.org/10.1109/T-C.1971.223208>.
- [54] Galle J, Hoffmann M, Aust G. From single cells to tissue architecture – a bottom-up approach to modelling the spatio-temporal organisation of complex multi-cellular systems. *J Math Biol* 2009;58(1–2):261–83. <https://doi.org/10.1007/s00285-008-0172-4>.
- [55] Garkavi LKh, Kvakina EB, Ukolova MA. *Adaptive reactions and resistance of the organism*. Rostov-on-Don: Rostov University Press; 1979.
- [56] Gause GF. *The struggle for existence*. Baltimore: Williams and Wilkins; 1934. Available from: <http://www.ggausa.com/Contgau.htm>.
- [57] Goldstone B. The general practitioner and the general adaptation syndrome. *S Afr Med J* 1952;26:88–92. PMID: 14901129. Available from: <http://archive.samj.org.za/1952%20VOL%20XXVI%20Jan-Jun/Articles/02%20February/1.5%20THE%20GENERAL%20PRACTITIONER%20AND%20THE%20GENERAL%20ADAPTION%20SYNDROME.%20Dr.%20B.%20Goldstone.pdf>.
- [58] Goldstone B. The general practitioner and the general adaptation syndrome. Concluded. *S Afr Med J* 1952;26:106–9. PMID: 14913266. Available from: <http://archive.samj.org.za/1952%20VOL%20XXVI%20Jan-Jun/Articles/02%20February/2.5%20THE%20GENERAL%20PRACTITIONER%20AND%20THE%20GENERAL%20ADAPTION%20SYNDROM%20-CONLUDED.%20Dr.%20B.%20Goldstone.pdf>.
- [59] Gomez LCM, Kondratova M, Sompairac N, Lonjou C, Ravel J-M, Barillot E, et al. Atlas of cancer signaling network: a resource of multi-scale biological maps to study disease mechanisms. In: Wolkenhauer O, editor. *Systems medicine: integrative qualitative and computational approaches. Reference module in biomedical sciences*. Elsevier; 2020.
- [60] Gorban AN. Selection theorem for systems with inheritance. *Math Model Nat Phenom* 2007;2(4):1–45.
- [61] Gorban AN. *Slow relaxations and bifurcations of omega-limit sets of dynamical systems*. PhD Thesis in Physics & Math (Differential Equations & Math.Phys), Kuibyshev, Russia. 1980. English translation was partially published in [62].
- [62] Gorban AN. Singularities of transition processes in dynamical systems: qualitative theory of critical delays. *Electron J Differ Equ Monogr* 2004;05. Available from: <https://ejde.math.txstate.edu/Monographs/05/gorban.pdf>. <https://arxiv.org/abs/chao-dyn/9703010>.
- [63] Gorban AN. Singularities of transient processes in dynamics and beyond: comment on “long transients in ecology: theory and applications” by Andrew Morozov et al. *Phys Life Rev* 2020;32:46–9. <https://doi.org/10.1016/j.plrev.2019.12.002>.
- [64] Gorban AN, Cheresiz VM. Slow relaxations of dynamical systems and bifurcations of ω -limit sets. *Dokl Akad Nauk SSSR* 1981;261(5):1050–3. communicated by S.L. Sobolev. English translation: *Sov Math Dokl* 1981;24:645–9.
- [65] Gorban AN, Golubkov A, Grechuk B, Mirkes EM, Tyukin IY. Correction of AI systems by linear discriminants: probabilistic foundations. *Inf Sci* 2018;466:303–22. <https://doi.org/10.1016/j.ins.2018.07.040>.
- [66] Gorban AN, Karlin IV. *Invariant manifolds for physical and chemical kinetics. Lecture notes in physics*, vol. 660. Springer; 2005.
- [67] Gorban AN, Kégl B, Wunsch D, Zinovyev A, editors. *Principal manifolds for data visualisation and dimension reduction. Lecture notes in computational science and engineering*, vol. 58. Berlin – Heidelberg – New York: Springer; 2008.
- [68] Gorban AN, Manchuk VT, Petushkova Smirnova EV. Dynamics of physiological parameters correlations and the ecological-evolutionary principle of polyfactoriality. *Problemy ekologicheskogo monitoringa I modelirovaniya ekosistem (The Problems of Ecological Monitoring and Ecosystem Modelling)*, vol. 10. Leningrad: Gidrometeoizdat; 1987. p. 187–98.
- [69] Gorban AN, Pokidysheva LI, Smirnova EV, Tyukina TA. Law of the minimum paradoxes. *Bull Math Biol* 2011;73(9):2013–44. <https://doi.org/10.1007/s11538-010-9597-1>.

- [70] Gorban AN, Smirnova EV, Tyukina TA. Correlations, risk and crisis: from physiology to finance. *Physica A* 2010;389(16):3193–217. <https://doi.org/10.1016/j.physa.2010.03.035>.
- [71] Gorban AN, Tyukin I, Prokhorov D, Sofeikov K. Approximation with random bases: pro et contra. *Inf Sci* 2016;364–365:129–45. <https://doi.org/10.1016/j.ins.2015.09.021>.
- [72] Gorban AN, Tyukina TA, Smirnova EV, Pokidysheva LI. Evolution of adaptation mechanisms: adaptation energy, stress, and oscillating death. *J Theor Biol* 2016;405:127–39. <https://doi.org/10.1016/j.jtbi.2015.12.017>.
- [73] Gorban AN, Zinovyev AY. Principal graphs and manifolds. Ch. 2. In: Soria Olivas Emilio, et al., editors. *Handbook of research on machine learning applications and trends: algorithms, methods, and techniques*. Hershey, PA: IGI Global; 2009. p. 28–59. Available from: <https://arxiv.org/abs/0809.0490>.
- [74] Gudayol-Ferré E, Guàrdia-Olmos J, Peró-Cebollero M. Effects of remission speed and improvement of cognitive functions of depressed patients. *Psychiatry Res* 2015;226(1):103–12. <https://doi.org/10.1016/j.psychres.2014.11.079>.
- [75] Giuliani A. Statistical mechanics of gene expression networks: increasing connectivity as a response to stressful condition. *Adv Syst Biol* 2014;3:1–4.
- [76] Haldane JBS. *The causes of evolution*. London: Longmans Green; 1932.
- [77] Harpole WS, Ngai JT, Cleland EE, Seabloom EW, Borer ET, Bracken MES, et al. Nutrient co-limitation of primary producer communities. *Ecol Lett* 2011;14(9):852–62. <https://doi.org/10.1111/j.1461-0248.2011.01651.x>.
- [78] Hessen DO, Anderson TR. Excess carbon in aquatic organisms and ecosystems: physiological, ecological, and evolutionary implications. *Limnol Oceanogr* 2008;53:1685–96. <https://doi.org/10.4319/lo.2008.53.4.1685>.
- [79] Hester RL, Iliescu R, Summers R, Coleman TG. Systems biology and integrative physiological modelling. *J Physiol* 2011;589(5):1053–60. <https://doi.org/10.1113/jphysiol.2010.201558>.
- [80] Honda M, Kaneko S, Kawai H, Shirota Y, Kobayashi K. Differential gene expression between chronic hepatitis B and C hepatic lesion. *Gastroenterology* 2001;120:955–66. <https://doi.org/10.1053/gast.2001.22468>.
- [81] Horn HS. Adaptation from the perspective of optimality. In: *Topics in plant population biology*. London: Palgrave; 1979. p. 48–61.
- [82] Hyvärinen A, Karhunen J, Oja E. *Independent component analysis*. New York: Wiley; 2001.
- [83] Huynh X-H, Guillet F, Briand H. Evaluating interestingness measures with linear correlation graph. In: *Advances in applied artificial intelligence. Lecture notes in computer science*, vol. 4031. Berlin - Heidelberg: Springer; 2006. p. 312–21.
- [84] Jolliffe H. *Principal component analysis*. 2nd ed. Springer series in statistics. New York: Springer; 2002.
- [85] Karanasos M, Yfanti S, Karoglou M. Multivariate FIAPARCH modelling of financial markets with dynamic correlations in times of crisis. *Int Rev Financ Anal* 2016;45:332–49. <https://doi.org/10.1016/j.irfa.2014.09.002>.
- [86] Karev GP, Kareva IG. Replicator equations and models of biological populations and communities. *Math Model Nat Phenom* 2014;9(3):68–95. <https://doi.org/10.1051/mmnp/20149305>.
- [87] Karmanova IV, Razzhevaikin VN, Shpitenkov MI. Application of correlation adaptometry for estimating a response of herbaceous species to stress loadings. *Dokl Bot Sci* 1996;346–348:4–7. Translated from *Dokl Akad Nauk SSSR* 1996;346.
- [88] Kleinbub JR, Palmieri A, Orsucci FF, Andreassi S, Musmeci N, Benelli E, et al. Measuring empathy: a statistical physics grounded approach. *Physica A* 2019;526:120979. <https://doi.org/10.1016/j.physa.2019.04.215>.
- [89] Kobe RK. Intraspecific variation in sapling mortality and growth predicts geographic variation in forest composition. *Ecol Monogr* 1996;66(2):181–201. <https://doi.org/10.2307/2963474>.
- [90] Kohl P, Crampin EJ, Quinn TA, Noble D. Systems biology: an approach. *Clin Pharmacol Ther* 2010;88(1):25–33. <https://doi.org/10.1038/clpt.2010.92>.
- [91] Koole SL, Tschacher W. Synchrony in psychotherapy: a review and an integrative framework for the therapeutic alliance. *Front Psychol* 2016;7:862. <https://doi.org/10.3389/fpsyg.2016.00862>.
- [92] Kuznetsov YuA. *Elements of applied bifurcation theory*. 3rd edition. Springer; 2004.
- [93] Lähnemann D, Köster J, Szczurek E, Schönhuth A. Eleven grand challenges in single-cell data science. *Genome Biol* 2020;21:31. <https://doi.org/10.1186/s13059-020-1926-6>.
- [94] Lan G, Sartori P, Neumann S, Sourjik V, Tu Y. The energy-speed-accuracy trade-off in sensory adaptation. *Nat Phys* 2012;8(5):422–8. <https://doi.org/10.1038/nphys2276>.
- [95] Landau LD. On the problem of turbulence. *Dokl Akad Nauk SSSR* 1944;44:339–42. Engl. transl. in *C R Acad Sci URSS* 1944;44:311–4. and in *Collected papers*. Oxford/New York: Pergamon Press/Gordon and Breach; 1965. p. 387–91.
- [96] Leo Y, Karsai M, Sarraute C, Fleury E. Correlations and dynamics of consumption patterns in social-economic networks. *Soc Netw Anal Min* 2018;8(1):9. <https://doi.org/10.1007/s13278-018-0486-1>.
- [97] Lewontin RC. *The genetic basis of evolutionary change*. New York: Columbia University Press; 1974.
- [98] Lion S. Theoretical approaches in evolutionary ecology: environmental feedback as a unifying perspective. *Am Nat* 2018;191(1):21–44. <https://doi.org/10.1086/694865>.
- [99] Longin F, Solnik B. Is the correlation in international equity returns constant: 1960–1990?. *J Int Money Financ* 1995;14(1):3–26. [https://doi.org/10.1016/0261-5606\(94\)00001-H](https://doi.org/10.1016/0261-5606(94)00001-H).
- [100] Loscalzo J, Barabasi AL, Silverman EK, editors. *Network medicine: complex systems in human disease and therapeutics*. Cambridge, MA: Harvard University Press; 2017.
- [101] Luo P, Shu K, Wu J, Wan L, Tan Y. Exploring correlation network for cheating detection. *ACM Trans Intell Syst Technol* 2020;11(1):1–23. <https://doi.org/10.1145/3364221>.
- [102] MacArthur RH, Wilson EO. *The theory of island biogeography*. Princeton, N.J.: Princeton Univ. Press; 1967.
- [103] Mansurov AS, Mansurova TP, Smirnova EV, Mikitin LS, Pershin AV. Human adaptation under influence of synergic system of factors (treatment of oncological patients after operation). Preprint 212B. Krasnoyarsk: Institute of Biophysics, Russian Academy of Sciences; 1995.

- [104] Marchenko VA, Pastur LA. Distribution of eigenvalues for some sets of random matrices. *Sb Math* 1967;114(4):507–36. <https://doi.org/10.1070/SM1967v001n04ABEH001994>.
- [105] Masaev SN, Dorrer MG. Company management system estimation on the basis of adaptive correlation to the environment. *Sib J Sci Technol* 2010;4(30):90–4. Available from: <http://mi.mathnet.ru/pu32>.
- [106] Masquelier T, Deco G. Network bursting dynamics in excitatory cortical neuron cultures results from the combination of different adaptive mechanisms. *PLoS ONE* 2013;8(10):e75824. <https://doi.org/10.1371/journal.pone.0075824>.
- [107] Mantegna RN, Stanley HE. *An introduction to econophysics: correlations and complexity in finance*. Cambridge: Cambridge University Press; 1999.
- [108] Maynard-Smith J. *Evolution and the theory of games*. Cambridge: Cambridge University Press; 1982.
- [109] McCarty R, Pasak K. Alarm phase and general adaptation syndrome. In: Fink George, editor. *Encyclopedia of stress*, vol. 1. Academic Press; 2000. p. 126–30.
- [110] Meric I, Meric G. Co-movements of European equity markets before and after the 1987 crash. *Multinational Finance J* 1997;1(2):137–52. <https://doi.org/10.17578/1-2-4>.
- [111] Metz JAJ, Nisbet RM, Geritz SAH. How should we define fitness for general ecological scenarios. *Trends Ecol Evol* 1992;7:198–202. [https://doi.org/10.1016/0169-5347\(92\)90073-K](https://doi.org/10.1016/0169-5347(92)90073-K).
- [112] Millington T, Niranjani M. Partial correlation financial networks. *Appl Netw Sci* 2020;5:11. <https://doi.org/10.1007/s41109-020-0251-z>.
- [113] Mirkes EM, Allohifi J, Gorban AN. Fractional norms and quasinorms do not help to overcome the curse of dimensionality. arXiv:2004.14230 [cs.LG], 2020. Available from: <https://arxiv.org/abs/2004.14230>.
- [114] Mirkes EM, Coats TJ, Levesley J, Gorban AN. Handling missing data in large healthcare dataset: a case study of unknown trauma outcomes. *Comput Biol Med* 2016;75:203–16. <https://doi.org/10.1016/j.combiomed.2016.06.004>.
- [115] Mitteroecker P, Bookstein F. The conceptual and statistical relationship between modularity and morphological integration. *Syst Biol* 2007;56(5):818–36. <https://doi.org/10.1080/10635150701648029>.
- [116] Neaime S. The global financial crisis, financial linkages and correlations in returns and volatilities in emerging MENA stock markets. *Emerg Mark Rev* 2012;13(3):268–82.
- [117] Neretti N, Remondini D, Tatar M, Sedivy JM, Pierini M, Mazzatti D. Correlation analysis reveals the emergence of coherence in the gene expression dynamics following system perturbation. *BMC Bioinform* 2007;8:S16. <https://doi.org/10.1186/1471-2105-8-S1-S16>.
- [118] Nica A, Speicher R. *Lectures on the combinatorics of free probability theory*. Cambridge Univ. Press; 2006. p. 204, 368.
- [119] Nijland GO, Schouls J, Goudriaan J. Integrating the production functions of Liebig, Michaelis-Menten, Mitscherlich and Liebscher into one system dynamics model. *NJAS Wageningen J Life Sci* 2008;55(2):199–224. [https://doi.org/10.1016/S1573-5214\(08\)80037-1](https://doi.org/10.1016/S1573-5214(08)80037-1).
- [120] Oechssler J, Riedel F. On the dynamic foundation of evolutionary stability in continuous models. *J Econ Theory* 2002;107:223–52. <https://doi.org/10.1006/jeth.2001.2950>.
- [121] Onnela J-P, Chakraborti A, Kaski K, Kertész J. Dynamic asset trees and the black monday. *Physica A* 2003;324:247–52. [https://doi.org/10.1016/S0378-4371\(02\)01882-4](https://doi.org/10.1016/S0378-4371(02)01882-4).
- [122] Onnela J-P, Kaski K, Kertész J. Clustering and information in correlation based financial networks. *Eur Phys J B* 2004;38:353–62. <https://doi.org/10.1140/epjb/e2004-00128-7>.
- [123] Orzack SH, Sober E. Optimality models and the test of adaptationism. *Am Nat* 1994;143(3):361–80. Available from: <https://www.jstor.org/stable/2462735>.
- [124] Pajor A, Pastur L. On the limiting empirical measure of eigenvalues of the sum of rank one matrices with log-concave distribution. *Stud Math* 2009;195:11–29. <https://doi.org/10.4064/sm195-1-2>.
- [125] Parker GA, Maynard Smith J. Optimality theory in evolutionary biology. *Nature* 1990;348(6296):27–33.
- [126] Pedronette DCG, Torres RDS. A correlation graph approach for unsupervised manifold learning in image retrieval tasks. *Neurocomputing* 2016;208:66–79.
- [127] Pescatori M, Ricci E, Bernardini C, Tonali PA. Expression data from quadriceps muscle of young DMD patients and age matched controls, gene expression omnibus (GEO). GEO Accession GSE6011. 2006. Available from: <http://www.ncbi.nlm.nih.gov/geo/>.
- [128] Pescatori M, Broccolini A, Minetti C, Bertini E, Bruno C, D’amico A, et al. Gene expression profiling in the early phases of DMD: a constant molecular signature characterizes DMD muscle from early postnatal life throughout disease progression. *FASEB J* 2007;21(4):1210–26. <https://doi.org/10.1096/fj.06-7285com>.
- [129] Pianka ER. On r - and K -selection. *Am Nat* 1970;104(940):592–7. Available from: <https://www.jstor.org/stable/2459020>.
- [130] Plerou V, Gopikrishnan P, Rosenow B, Amaral LAN, Guhr T, Stanley HE. Random matrix approach to cross correlations in financial data. *Phys Rev E* 2002;65:066126. <https://doi.org/10.1103/PhysRevE.65.066126>.
- [131] Pokidysheva EV. An estimation of conjugacy of monetary and banking policy during crisis. *Finance Credit* 2010;16(42):72–8. Available from: <https://www.fin-izdat.com/journal/fc/detail.php?ID=32620>.
- [132] Ponomarenko LD, Smirnova EV. Dynamical characteristics of blood system in mice with phenylhydrazin anemiya. In: *Proc. of 9th international symposium “reconstruction of homeostasis”*, vol. 1; 1998. p. 42–5.
- [133] Potters M, Bouchaud JP, Laloux L. Financial applications of random matrix theory: old laces and new pieces. *Acta Phys Pol B* 2005;36(9):2767–84. Available from: <https://arxiv.org/abs/physics/0507111>.
- [134] Pough FH. Performance and Darwinian fitness: approaches and interpretations. *Physiol Zool* 1989;62(2):199–236. <https://doi.org/10.1086/physzool.62.2.30156169>.
- [135] Razzhevaikin VN, Shpitonkov MI. Substantiation of correlation adaptometry based on evolutionary optimality principles. *Comput Math Math Phys* 2003;43(2):296–307. Available from: <http://mi.mathnet.ru/zvmmf1068>.
- [136] Razzhevaikin VN, Shpitonkov MI. Correlation adaptometry. Models and applications to biomedical systems. *Mat Model* 2008;20(8):13–27. Available from: <http://mi.mathnet.ru/mm2671>.

- [137] Red'ko VG, Prokhorov DV, Burtsev MS. Theory of functional systems, adaptive critics and neural networks. In: Proc. 2004 IEEE international joint conference on neural networks (IEEE Cat. No. 04CH37541), vol. 3. IEEE; 2004. p. 1787–92.
- [138] Reilly T, Doran D. Fitness assessment. In: Reilly T, Williams AM, editors. Science and soccer. New York: The Taylor & Francis; 2003. p. 26–46.
- [139] Rybnikov SR, Rybnikova NA, Portnov BA. Public fears in Ukrainian society: are crises predictable?. Psychol Dev Soc 2017;29(1):98–123. <https://doi.org/10.1177/0971333616689398>.
- [140] Salisbury F. Plant physiology. 4th ed. Belmont, CA: Wadsworth; 1992.
- [141] Saito MA, Goepfert TJ. Some thoughts on the concept of colimitation: three definitions and the importance of bioavailability. Limnol Oceanogr 2008;53(1):276–90. <https://doi.org/10.4319/lo.2008.53.1.0276>.
- [142] Scheffer M, Carpenter SR, Lenton TM, Bascompte J, Brock W, Dakos V, et al. Anticipating critical transitions. Science 2012;338(6105):344–8. <https://doi.org/10.1126/science.1225244>.
- [143] Schkade JK, Schultz S. Occupational adaptation. In: Kramer Paula, Hinojosa Jim, Royeen Charlotte Brasic, editors. Perspectives in human occupation: participation in life. Baltimore, MD: Lippincott Williams & Wilkins; 2003. p. 181–221.
- [144] Schneck DJ. Feedback control and the concept of homeostasis. Math Model 1987;9(12):889–900. [https://doi.org/10.1016/0270-0255\(87\)90007-8](https://doi.org/10.1016/0270-0255(87)90007-8).
- [145] Sciuto AM, Phillips CS, Orzolek LD, Hege AI, Moran TS, Dillman JF. Genomic analysis of murine pulmonary tissue following carbonyl chloride inhalation. Chem Res Toxicol 2005;18(11):1654–60. <https://doi.org/10.1021/tx050126f>.
- [146] Sedov KR, Gorban' AN, Petushkova Smirnova EV, Manchuk VT, Shalamova EN. Correlation adaptometry as a method of screening of the population. Vestn Akad Med Nauk SSSR 1988;(10):69–75. PMID: 3223045.
- [147] Selye H. Adaptation energy. Nature 1938;141(3577):926. <https://doi.org/10.1038/141926a0>.
- [148] Selye H. Experimental evidence supporting the conception of “adaptation energy”. Am J Physiol 1938;123:758–65. <https://doi.org/10.1152/ajplegacy.1938.123.3.758>.
- [149] Sengupta AM, Mitra PP. Distributions of singular values for some random matrices. Phys Rev E 1999;60:3389–92. <https://doi.org/10.1103/PhysRevE.60.3389>.
- [150] Shpitonkov MI. Application of correlation adaptometry technique to sports and biomedical research. Comput Res Model 2017;9(2):345–54. <https://doi.org/10.20537/2076-7633-2017-9-2-345-354>.
- [151] Shumeiko PG, Osipov VI, Kofman GB. Early detection of industrial emission impact on Scots Pine needles by composition of phenolic compounds. In: Khlebopros RG, editor. Global & regional ecological problems. Krasnoyarsk: Krasnoyarsk State Technical University Publ.; 1994. p. 536–43.
- [152] Sih A, Gleeson SK. A limits-oriented approach to evolutionary ecology. Trends Ecol Evol 1995;10(9):378–82. [https://doi.org/10.1016/S0169-5347\(00\)89142-9](https://doi.org/10.1016/S0169-5347(00)89142-9).
- [153] Silvennoinen A, Thorp S. Financialization crisis and commodity correlation dynamics. J Int Financ Mark 2013;24:42–65. <https://doi.org/10.1016/j.intfin.2012.11.007>.
- [154] Smith SL, Pahlow M, Merico A, Wirtz KW. Optimality-based modeling of planktonic organisms. Limnol Oceanogr 2011;56(6):2080–94. <https://doi.org/10.4319/lo.2011.56.6.2080>.
- [155] Soloviev VN, Belinskiy A. Complex systems theory and crashes of cryptocurrency market. In: Proc. international conference on information and communication technologies in education, research, and industrial applications ICTERI 2018. Cham: Springer; 2018. p. 276–97.
- [156] Sompairac N, Nazarov PV, Czerwinska U, Cantini L, Biton A, Molkenov A, et al. Component analysis for unraveling the complexity of cancer omics datasets. Int J Mol Sci 2019;20:4414. <https://doi.org/10.3390/ijms20184414>.
- [157] Sperfeld E, Martin-Creuzburg D, Wacker A. Multiple resource limitation theory applied to herbivorous consumers: Liebig's minimum rule vs. interactive co-limitation. Ecol Lett 2012;15(2):142–50. <https://doi.org/10.1111/j.1461-0248.2011.01719.x>.
- [158] Strygina SO, Dement'ev SN, Uskov VM, Chernyshova GI. Dynamics of the system of correlations between physiological parameters in patients after myocardial infarction. In: Proc. of conference mathematics, computer, education, issue 7; 2000. p. 685–9.
- [159] Svetlichnaya GN, Smirnova EV, Pokidysheva LI. Correlation adaptometry in studies of cardiovascular and respiratory relationships. Fiziol Cheloveka 1997;23(3):58–62. Fiziol Cheloveka 1997;23(3):319–23. PMID: 9264951.
- [160] Sudakov KV. Functional systems theory: a new approach to the question of the integration of physiological processes in the body. Neurosci Behav Physiol 2004;34(5):495–500.
- [161] Terentjev PV. Biometrische Untersuchungen Über Die Morpho-Logischen Merkmale Von Rana Ridibunda Pall: (Amphibia, Salientia). Biometrika 1931;1:23–51. Available from: <http://www.jstor.org/stable/2333629>.
- [162] Tilman D. Resources: a graphical-mechanistic approach to competition and predation. Am Nat 1980;116(3):362–93. Available from: <https://www.jstor.org/stable/2463311>.
- [163] Tilman D. Resource competition and community structure. Princeton, NJ: Princeton University Press; 1982.
- [164] Triviño V, Nuño de la Rosa L. A causal dispositional account of fitness. Hist Philos Life Sci 2016;38:6. <https://doi.org/10.1007/s40656-016-0102-5>.
- [165] Tyukin I. Adaptation in dynamical systems. NY: Cambridge University Press; 2011.
- [166] Tyukin IY, Iudin D, Iudin F, Tyukina T, Kazantsev V, Mukhina I, et al. Simple model of complex dynamics of activity patterns in developing networks of neuronal cultures. PLoS ONE 2019;14(6):e0218304. <https://doi.org/10.1371/journal.pone.0218304>.
- [167] Utsugi A, Ino K, Oshikawa M. Random matrix theory analysis of cross correlations in financial markets. Phys Rev E 2004;70:026110. <https://doi.org/10.1103/PhysRevE.70.026110>.
- [168] van der Ploeg RR, Böhm W, Kirkham MB. On the origin of the theory of mineral nutrition of plants and the law of the minimum. Soil Sci Soc Am J 1999;63:1055–62. <https://doi.org/10.2136/sssaj1999.6351055x>.
- [169] Vasenina E, Kataoka R, Buckner SL. Adaptation energy: experimental evidence and applications in exercise science. J Trainology 2020;9(2):66–70. https://doi.org/10.17338/trainology.9.2_66.

- [170] Vasil'ev AV, Mal'tsev GIu, Khrushcheva IuV, Razzhevaikin VN, Shpitionkov MI. Applying method of correlation adaptometry for evaluating of treatment efficiency of obese patients. *Vopr Pitan* 2007;76(2):36–8. PMID: 17561653.
- [171] Verma V, Gagvani N. Visualizing intelligence information using correlation graphs. *Proc SPIE* 2005;5812:271–82. <https://doi.org/10.1117/12.604040>.
- [172] Wagner DE, Klein AM. Lineage tracing meets single-cell omics: opportunities and challenges. *Nat Rev Genet* 2020;21:410–27. <https://doi.org/10.1038/s41576-020-0223-2>.
- [173] Wang C, Bi J, Rikkert MGO. Early warning signals for critical transitions in cardiopulmonary health, related to air pollution in an urban Chinese population. *Environ Int* 2018;121:240–9. <https://doi.org/10.1016/j.envint.2018.09.007>.
- [174] Wang GJ, Xie C, Lin M, Stanley HE. Stock market contagion during the global financial crisis: a multiscale approach. *Finance Res Lett* 2017;22:163–8. <https://doi.org/10.1016/j.frl.2016.12.025>.
- [175] Wang Y, Ding S, Xu X, Jia W. The multi-tag semantic correlation used for micro-blog user interest modeling. *Eng Appl Artif Intell* 2019;85:765–72. <https://doi.org/10.1016/j.engappai.2019.08.007>.
- [176] Wainwright PC. Functional morphology as a tool for ecological research. In: Wainwright PC, Reilly SM, editors. *Ecological morphology: integrative organismal biology*. Chicago: University of Chicago Press; 1994. p. 42–59.
- [177] Whittaker J. *Graphical models in applied multivariate statistics*. Chichester: Wiley; 1990.
- [178] Wishart J. The generalised product moment distribution in samples from a normal multivariate population. *Biometrika* 1928;20A(1–2):32–52. <https://doi.org/10.1093/biomet/20A.1-2.32>.
- [179] Woldeamay AA, Ntwasa M. Pathways and network based analysis of candidate genes to reveal cross-talk and specificity in the sorghum (*Sorghum bicolor* (L.) Moench) responses to drought and it's co-occurring stresses. *Front Genet* 2018;9:557. <https://doi.org/10.3389/fgene.2018.00557>.
- [180] Xu R, Wunsch D. *Clustering*. Hoboken, New Jersey: John Wiley & Sons Inc.; 2008.
- [181] Zhang W, Wahl LM, Yu P. Modeling and analysis of recurrent autoimmune disease. *SIAM J Appl Math* 2014;74(6):1998–2025. <https://doi.org/10.1137/140955823>.
- [182] Zimatore G, Garilli G, Poscolieri M, Rafanelli C, Gizzi FT, Lazzari M. The remarkable coherence between two Italian far away recording stations points to a role of acoustic emissions from crustal rocks for earthquake analysis. *Chaos* 2017;27(4):043101. <https://doi.org/10.1063/1.4979351>.
- [183] Zinn M, Witholt B, Egli T. Dual nutrient limited growth: models, experimental observations, and applications. *J Biotechnol* 2004;113(1–3):263–79. <https://doi.org/10.1016/j.jbiotec.2004.03.030>.

Comment

Searching for unifying laws of general adaptation syndrome

Comment on “Dynamic and thermodynamic models of adaptation” by Gorban et al.

Kumar Selvarajoo ^{a,b,c,*}^a Bioinformatics Institute (BII), Agency for Science, Technology and Research (A*STAR), Biopolis, Singapore^b Singapore Institute of Food and Biotechnology Innovation (SIFBI), Agency for Science, Technology and Research (A*STAR), Biopolis, Singapore^c Synthetic Biology for Clinical and Technological Innovation (SynCTI), National University of Singapore (NUS), Medical Drive, Singapore

Received 6 April 2021; accepted 6 April 2021

Available online 7 April 2021

Communicated by L. Perlovsky

Non-linear systems, such as the living organisms, are remarkably adaptable for a wide variety of perturbations or drastic environmental changes. Notably, certain responses are spontaneous and are not controllable by the conscious mind. For example, when we touch an hot object like a cooking pan, we shout and remove our body part instantaneously even before our mind help realize the next steps, e.g. switching-off gas hob, apply soothing cream, etc. Although this stress response is a common household example, nevertheless, it makes us realize that our body parts, tissues, cells and molecular species are highly interconnected and tightly regulated for a survival outcome.

The review by Gorban and colleagues [1] highlights a relatively old concept on adaptation energy initially proposed by Hans Selye in 1938 [2,3], that consists of 3 parts; alarm, resistance and exhaustion. This concept is also referred to as the general adaptation syndrome (GAS) [4]. The basic premise is that when a body detects danger, it initiates a biological process called a fight-or-flight response. If the stress condition persists, the body reacts metabolically to control or neutralize the stress. On the longer term, the body will be unable to resist the stress as its resources will be lost, subsequently leading to serious health issues or even death. It is, thus, conceivable that understanding GAS mechanistically, and finding predictive models will help prevent illnesses and save lives.

Now 83 years after the initial concept was first formally introduced, there are multitude of evidences to substantiate the existence of adaptation energy or GAS. This is precisely what Gorban and colleagues have elegantly put across their entire review, succinctly covering, not only for living cells, but across multiple disciplines such as in physics, finance and social systems.

The review can be broadly classified into two parts. The first part of the review deals with simple *dynamic correlation* and *variance* studies of key biomarkers of disease states, as well as multi-dimensional datasets, e.g. high throughput gene expressions analytics. For all these works, it is stressed that the 2 statistical metrics are very useful markers for adaptation and stress and that they, collectively, relate to the 3 parts of GAS (Figs. 2 and 3 of [1]). The

DOI of original article: <https://doi.org/10.1016/j.plrev.2021.03.001>.

* Correspondence to: BII, A*STAR, Biopolis, Singapore.

E-mail address: kumar_selvarajoo@bii.a-star.edu.sg.

main findings from several highlighted works indicate that both correlations and variances tend to increase for initial stages of stress or illness. At the exhaustion stage or near death, the correlation may decrease, however, the variances can increase [5]. The authors noted that this is shown only when the relevant attributes are correctly selected. Specifically, it is cautioned that for this later part (exhaustion or death) the available data are still ambiguous and require further experimentations, and new theories need to be explored.

To overcome the limitations of simple statistics, the second half deals with the adaption thermodynamics. Here, Gorban and colleagues have reviewed several dynamical systems models to observe whether there are any guiding principles that can eventually lead to the development of an unified theory for GAS. Having such theory or model will help predict complexity, such as the onset of illnesses and treatment strategies for controlling or preventing the illnesses or even death. They have reviewed several dynamic models including a kinetic-resource model [6], which is based on a so-called *middle-out* approach; a hybrid of bottom-up (which requires all systems details) and top-down (more phenomical or empirical) approaches [7]. Models like these are able to show emergent behaviors such as oscillations, multiple stable states or unstable regions [8], providing insights into non-intuitive or unexpected behaviors.

For the last two decades, our laboratory has worked on high-throughput gene expressions data using multidimensional analytics [9,10], and both deterministic and stochastic *middle-out* models [11–14], chiefly to interpret and predict the complex immune, cancer and biofilm responses. Overtly, I do agree with the detailed examples and the postulations made by Gorban and colleagues. However, there are several other factors that require mentioning or could be highlighted with some degree of discussion. One important factor is biological noise [15]. When studying large scale gene expressions, even between replicates data, gene expressions variability from cell to cell is observed [16]. This variation can be independent from technical noise, such as protocol and user variances or instrumental effects, and can be a hallmark for non-linear responses such as bifurcations [17].

For the first part of correlation and variance studies, it is unclear how the authors of the original papers treated technical and biological noises. Most often, experimentalists discard low expression genes, proteins or metabolites using stringent arbitrary threshold cutoff [18]. Although this may seem reasonable on the first impression, to observe statistical structures or features, this might not be a good choice, as eliminating large number of data species will affect the overall statistical metrics [18]. In addition, replicate data variability should also be compared to estimate the level of noise that each sample carries [19]. This becomes even more cumbersome for patient-to-patient data. Once such quality control checks have been performed or verified for the first part of correlation and variance studies, the conclusion can be more convincing. In addition, the authors suggest choosing the correct attributes will lead to robust results or conclusion. How can we determine the correct attributes and, what criteria should one use?

On the hand, for the dynamic models, especially for the non-linear ones, the reliable estimation of parameter values is crucial, as small changes can lead to drastic changes in final outcome. This is especially important for chaos models, and notably, biological systems are known to display the properties of chaos [17]. On top of the technical variability issue, we also need to consider the effects of cell-to-cell variability [16]. Addressing some of these important issues will persuade readers to consider GAS more seriously.

Nevertheless, and in overall, I am deeply satisfied with the general concepts presented on GAS and the numerous examples that Gorban and colleagues have painstaking studied and presented in a highly unified manner. This effort requires both acknowledgment and support moving forward. Most importantly, we are in a continuous search for defining guiding principles and laws that shape life [20,21], which may not be simple nor straightforward and may well be reminiscent to finding the laws of our universe. I believe this work will enlighten both experimentalists and theoreticians alike to explore the governing rules for GAS, especially those who still view living systems in a highly reductionist manner.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The author thank A*STAR for funding this work.

References

- [1] Gorban AN, Tyukina TA, Pokidysheva LI, Smirnova EV. Dynamic and thermodynamic models of adaptation. *Phys Life Rev* 2021;37:17–64. <https://doi.org/10.1016/j.plrev.2021.03.001> [in this issue].
- [2] Selye H. Adaptation energy. *Nature* 1938;141(3577):926.
- [3] Selye H. Experimental evidence supporting the conception of “adaptation energy”. *Am J Physiol* 1938;123:758–65.
- [4] Selye H. Stress and the general adaptation syndrome. *Br Med J* 1950;1(4667):1383–92.
- [5] Giuliani A. Statistical mechanics of gene expression networks: increasing connectivity as a response to stressful condition. *Adv Syst Biol* 2014;3:1–4.
- [6] Gorban AN, Tyukina TA, Smirnova EV, Pokidysheva LI. Evolution of adaptation mechanisms: adaptation energy, stress, and oscillating death. *J Theor Biol* 2016;405:127–39.
- [7] Edwards LM, Thiele I. Applying systems biology methods to the study of human physiology in extreme environments. *Extrem Physiol Med* 2013;2(1):8.
- [8] Selvarajoo K. Complexity of biochemical and genetic responses reduced using simple theoretical models. *Methods Mol Biol* 2018;1702:171–201.
- [9] Tsuchiya M, Piras V, Choi S, Akira S, Tomita M, Giuliani A, et al. Emergent genome-wide control in wildtype and genetically mutated lipopolysaccharides-stimulated macrophages. *PLoS ONE* 2009;4(3):e4905.
- [10] Simeoni O, Piras V, Tomita M, Selvarajoo K. Tracking global gene expression responses in T cell differentiation. *Gene* 2015;569(2):259–66.
- [11] Hayashi K, Piras V, Tabata S, Tomita M, Selvarajoo K. A systems biology approach to suppress TNF-induced proinflammatory gene expressions. *Cell Commun Signal* 2013;11:84.
- [12] Piras V, Hayashi K, Tomita M, Selvarajoo K. Investigation of stochasticity in TRAIL signaling cancer model. *Proc IEEE/ICME Com Med Eng* 2012:609–14.
- [13] Deveaux W, Hayashi K, Selvarajoo K. Defining rules for cancer cell proliferation in TRAIL stimulation. *NPJ Syst Biol Appl* 2019;5:5.
- [14] Deveaux W, Selvarajoo K. Searching for simple rules in *Pseudomonas aeruginosa* biofilm formation. *BMC Res Notes* 2019;12(1):763.
- [15] Eldar A, Elowitz MB. Functional roles for noise in genetic circuits. *Nature* 2010;467(7312):167–73.
- [16] Huang S. Non-genetic heterogeneity of cells in development: more than just noise. *Development* 2009;136(23):3853–62.
- [17] Selvarajoo K. Understanding multimodal biological decisions from single cell and population dynamics. *Wiley Interdiscip Rev, Syst Biol Med* 2012;4(4):385–99.
- [18] Dalman MR, Deeter A, Nimishakavi G, Duan ZH. Fold change and p-value cutoffs significantly alter microarray interpretations. *BMC Bioinform* 2012;13(Suppl 2):S11.
- [19] Bui TT, Lee D, Selvarajoo K. ScatLay: utilizing transcriptome-wide noise for identifying and visualizing differentially expressed genes. *Sci Rep* 2020;10(1):17483.
- [20] Selvarajoo K, Tomita M. Physical laws shape biology. *Science* Feb 2013;339(6120):646.
- [21] Bizzarri M, Naimark O, Nieto-Villar J, Fedeli V, Giuliani A. Complexity in biological organization: deconstruction (and subsequent restating) of key concepts. *Entropy (Basel)* 2020;22(8):885.

Comment

The statistical mechanics of life

Comment on “Dynamic and thermodynamic models of adaptation” by A.N. Gorban et al.

Alessandro Giuliani

Istituto Superiore di Sanità, Viale Regina Elena 299, 00161, Roma, Italy

Received 10 April 2021; accepted 12 April 2021

Available online 14 April 2021

Communicated by L. Perlovsky

The actual reproducibility crisis of biological sciences [1] asks for seriously facing the relational character of biological matter. This focus reshaping is advocated by many scientists and stems from the ‘Middle Way’ [2] approach, i.e. the mesoscopic description of complex systems, adopting as privileged level of observation the one that maximizes the correlation among variables relative to different scales of observation [3–5].

The mesoscopic approach unveils the presence of ‘organization principles’ whose main peculiarity is the independence from the detailed knowledge of the basic bricks of the system and the emergence of relevant features from the correlation structure of the studied phenomenon. This attitude shifts the ‘shared foundation of different sciences’ from the recognition that ‘*all the entities are made of the same fundamental particles*’ (orienting the various ‘theories of everything’ [6] flourished in the last century) to the statement ‘*all the entities can be considered as networks of interacting parts*’.

The adoption of this second viewpoint, shifts the frontiers of fundamental science from the world of high energy physics to the development of a ‘network thermodynamics’ [7]. At odds with classical thermodynamics, organized complexity cannot rely neither on basic invariance of interactions, nor on a huge number of substantially identical particles. As already stressed in 1948 by Warren Weaver [8], organized complexity shows relational structures changing in time by the action of both external stressors and development. This implies that the foundation of a biological statistical mechanics should focus on the changes in correlation patterns marking state transitions.

This is exactly the way the authors of the commented paper follow by aptly reporting a lot of examples coming from different investigation fields (by the way universality is a pillar of statistical mechanics and corresponds to the existence of ‘organization principles’ largely independent of, while consistent with constitutive microscopic laws).

The exploration of organized complexity involves a peculiar relation between quantitative tools and the investigated phenomena. Mathematical tools adopted in science can be partitioned into two general classes [9]. One class is made of methods stemming from some specific theory or tailored around a given phenomenon (e.g. renormalization group or S matrix formalism in theoretical physics, SIR models in epidemiology..). The other class combines under a single rubric all those methods designated as statistical. Such methods are independent of any specific application, and refer to *a posteriori* check of specific hypotheses on the system at hand, remaining ‘neutral’ with respect to any content

DOI of original article: <https://doi.org/10.1016/j.plrev.2021.03.001>.E-mail address: alessandro.giuliani@iss.it.

specific model. In general these methods come into play at the end of the experimentation to assign a likelihood estimate of the starting hypothesis. Examples of this class are Analysis of Variance or Student's t-test. Some rare cases lie in between these two extremes so blurring the neat distinction between general purpose statistical techniques and theoretical intensive methods. Principal Component Analysis (PCA) is the most widespread of these 'intermediate' quantitative tools: while being a general purpose statistical technique, it implies a peculiar style of reasoning. PCA is a 'hypothesis generating' tool creating a statistical mechanics frame for modeling without the need of strong *a priori* theoretical assumptions impossible to satisfy in the case of biology. PCA and related correlation techniques like the G index proposed in the commented paper, provide a sort of 'ready-made' theoretical framework by directly connecting (by the agency of the correlation matrix among the observed variables) the studied phenomenon to physical concepts like the degree of order [10] or the attractor state of a dynamical system [11]. This state of affairs, deeply influences the use of mathematics in scientific work: while in 'hard' sciences the mathematics 'precedes' the experiment setting theoretically motivating quantitative predictions to be eventually verified, in 'soft' sciences the hypothesis to be tested stems from a narrative consistent with previous knowledge and the mathematics come 'after' the experiment in the form of statistical analysis of the results. The field of 'organized complexity' combines these two extremes mixing up hypothesis generating and hypothesis testing phases: the use of mathematics goes together with the experimentation and mathematical methods work like laboratory instruments as microscopes or spectrophotometers.

This new role played by mathematical methods greatly enlarges the reach of phenomena amenable of a quantitative approach: the authors of the commented paper analyze the hypothesis of the existence of 'adaptation energy' set forth in the thirties by Hans Selye [12]. The concept of 'adaptation energy' stems from the common empirical experience of the increased sensitivity to other diseases of individuals experiencing (and consequently 'stressed by') a given pathology.

The idea of 'adaptation energy' was dismissed by scientific community for the inability to find a material counterpart to this energy, even if the increased vulnerability to other stressors of people affected by a specific disease is a shared tenet of any medical practitioner (the striking evidence of the increased lethality of COVID-19 for people bearing other pathologies is a clear example of the crucial role played by adaptation energy).

In their work, the authors are able to give a 'quantitative' counterpart to the adaptation energy in terms of correlation dynamics. If this represents a 'material' counterpart of adaptation energy is, at least in my opinion, a largely irrelevant issue that can be equated to the discussion if electromagnetic field is material or not.

The enlargement of the spectrum of the phenomena amenable of a quantitative approach is in my opinion the most important (and revolutionary) contribute of this paper: in the actual pandemics we experienced the apparently insurmountable divide between largely empirical physicians wisdom and laboratory science. The promise of 'closing the gap' set forth by ultra-reductionist gene-centric approach largely failed in this enterprise [13] on the other hand, purely computational Big Data approaches can only offer empirical solutions to specific problems; the development of a 'Statistical Mechanics of Life' could be the beginning of a new alliance.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] Ioannidis JPA. Why most published research findings are false. *PLoS Med* 2005;2(8):e124.
- [2] Laughlin RB, Pines D, Schmalian J, Stojković BP, Wolynes P. The middle way. *Proc Natl Acad Sci USA* 2000;97(1):32–7.
- [3] Hastings A, Petrovskii S, Morozov A. Spatial ecology across scales. *Biol Lett* 2011;75:163–5.
- [4] Pascual M, Levin SA. From individuals to population densities: searching for the intermediate scale of nontrivial determinism. *Ecology* 1999;80(7):2225–36.
- [5] Noble R, Tasaki K, Noble PJ, Noble D. Biological relativity requires circular causality but not symmetry of causation: so, where, what and when are the boundaries? *Front Physiol* 2019;10:827.
- [6] Laughlin RB, Pines D. The theory of everything. *Proc Natl Acad Sci USA* 2000;97(1):28–31.
- [7] Mikulecki DC. Network thermodynamics and complexity: a transition to relational systems theory. *Comput Chem* 2001;25(4):369–91.
- [8] Weaver W. Science and complexity. *Am Sci* 1948;36:536–41.
- [9] Webber Jr CL, Marwan N, Facchini A, Giuliani A. Simpler methods do it better: success of recurrence quantification analysis as a general purpose data analysis tool. *Phys Lett A* 2009;373(41):3753–6.

- [10] Giuliani A. The application of principal component analysis to drug discovery and biomedical data. *Drug Discov Today* 2017;22(7):1069–76.
- [11] Broomhead DS, King GP. Extracting qualitative dynamics from experimental data. *Phys D Nonlinear Phenom* 1986;20(2–3):217–36.
- [12] Selye H. Adaptation energy. *Nature* 1938;141(3577):926.
- [13] Voosen P. Amid a sea of false findings, the NIH tries reform. *Chron High Educ* 2015:228479.

Comment

Collective fluctuation implies imminent state transition

Comment on “Dynamic and thermodynamic models of adaptation” by A.N. Gorban et al.

Rui Liu^{a,b}, Kazuyuki Aihara^{c,d,*}, Luonan Chen^{e,f,g,*}^a School of Mathematics, South China University of Technology, Guangzhou 510640, China^b Pazhou Lab, Guangzhou 510330, China^c International Research Center for Neurointelligence, The University of Tokyo, Tokyo 113-8654, Japan^d Institute of Industrial Science, The University of Tokyo, Tokyo 153-8505, Japan^e State Key Laboratory of Cell Biology, Shanghai Institute of Biochemistry and Cell Biology, Center for Excellence in Molecular Cell Science, Chinese Academy of Sciences, Shanghai 200031, China^f Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming 650223, China^g Key Laboratory of Systems Health Science of Zhejiang Province, Hangzhou Institute for Advanced Study, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Hangzhou 310024, China

Received 10 April 2021; accepted 12 April 2021

Available online 19 April 2021

Communicated by L. Perlovsky

Critical transitions or tipping points widely exist in many natural and social systems including biological systems. Identifying the tipping point or the critical state of a complex disease before disease onset or symptom appearance is of great importance for disease prevention and early treatment. However, in contrast to the disease state, identifying the critical state is a difficult task due to its similarities to the normal state in terms of phenotypes and molecular expressions, which make traditional static biomarkers fail. A complex disease is generally caused not from the malfunction of individual molecules but from the interplay of a group of correlated molecules or a network [25]. With the background of dynamical systems theory including the bifurcation and critical slowing down [28], a concept called dynamical network biomarkers (DNB) was developed to quantitatively characterize the critical state during the disease progression [2]. Specifically, the development of a disease is considered as a time-dependent nonlinear dynamical system with a tipping point through which the system falls into a disease state [18], which can be mathematically described as the state/phase transition through a bifurcation at the tipping point (Fig. 1a), as shown from the well-written review by Gorban et al. [6]. Thus, when a dynamical system approaches a critical state, a dominant group often appears and can be defined as the DNB, i.e. a dynamical network of biomarkers, which meets the following two necessary conditions in terms of the observed data with noise. (i) The coefficient of variation for any member/variable in the DNB group rapidly increases. (ii) The Pearson's correlation coefficients in absolute values between the DNB members/variables rapidly increase [2,8,21]. This theoretical result means that appearance of a group of collectively fluctuating variables among the observed high-dimensional data at a healthy state signals an imminent critical transition, or briefly, strongly collective fluctuation implies an imminent state transition from a healthy state to a disease state. Being consistent with

DOI of original article: <https://doi.org/10.1016/j.plrev.2021.03.001>.

* Corresponding author.

E-mail addresses: scliurui@scut.edu.cn (R. Liu), kaihara@g.ecc.u-tokyo.ac.jp (K. Aihara), lnchen@sibs.ac.cn (L. Chen).

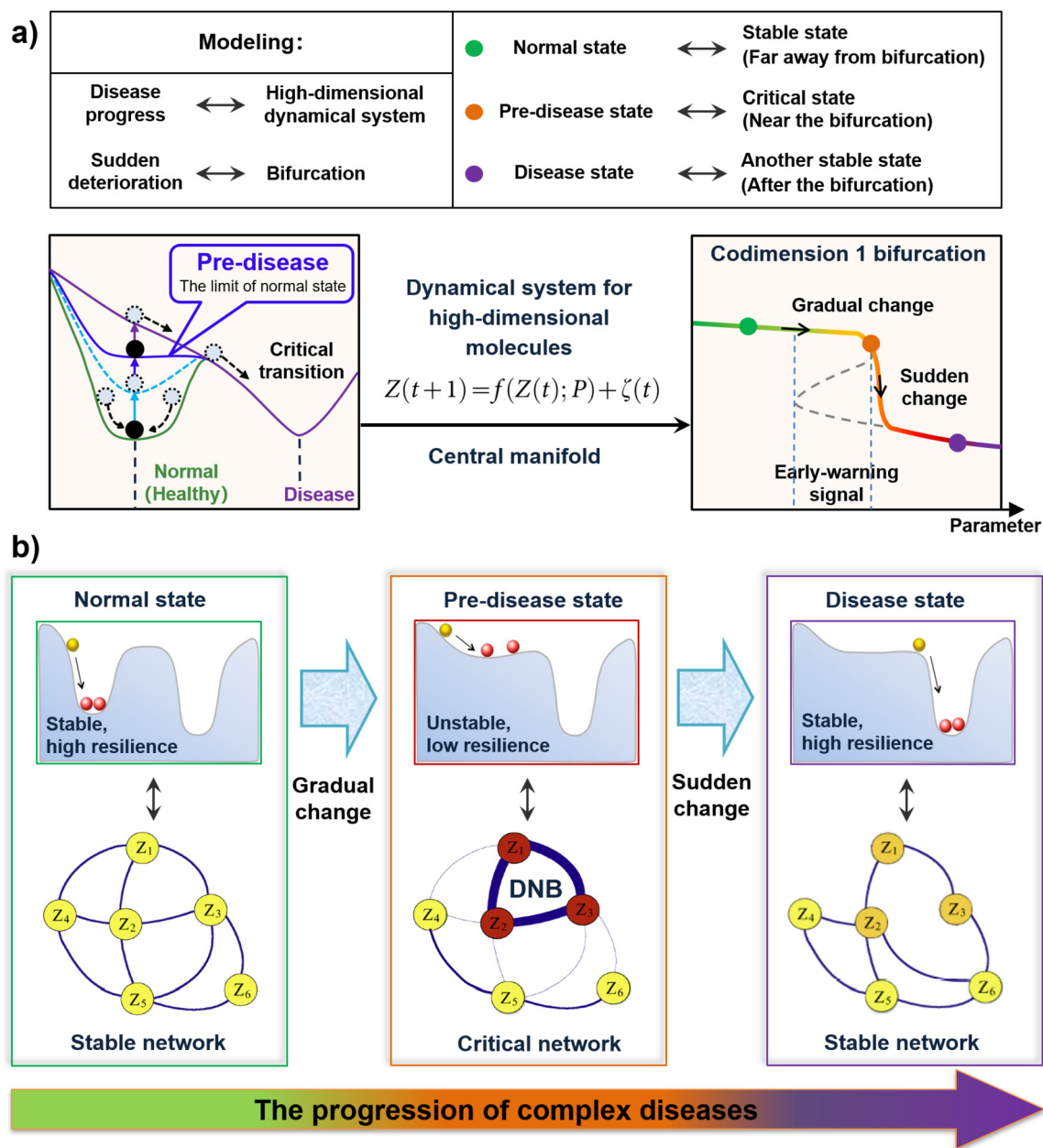


Fig. 1. Schematic illustration of dynamical network biomarker (DNB). **(a)** The disease progression can be viewed as the evolution of a nonlinear dynamical system with steady states and also a bifurcation/tipping point. Generally, the disease progression can be divided into three stages or states, i.e., a normal state, a pre-disease/critical state, and a disease state. **(b)** When the system approaches the pre-disease state with a low resilience, the critical subnetwork or DNB appears, which contains a group of biomolecules with wild fluctuations in a strong collective manner, i.e. collective fluctuation implies an imminent state transition.

the correlation graph presented by Gorban et al. [6] in terms of correlation and variance, such two generic properties can be used to quantify the instability or transition of a dynamical system in terms of network changes (Fig. 1b), that is, after the gradual changes in the system, a subnetwork or DNB appears, and its members strongly fluctuate in a collective manner, thus implying the presence of an imminent state transition. In addition to the above two conditions, the following third condition is also sometimes used: (iii) The Pearson's correlation coefficients in absolute values between the DNB and non DNB members/variables rapidly decrease.

The DNB method has been applied to identify critical states or pre-disease states of a number of chronic diseases including liver cancer, lymphoma cancer, lung cancer and type-2 diabetes as well as acute diseases based on the omics data [5,12,13,17,27,29]. Moreover, since the proposal of such a concept, the related methods and its follow-up modifications have been employed by researchers for analysis of biological data and other type of disease data, e.g., identifying the tipping points of drug resistance [17], detecting the tipping points of cell fate decision [23] as well as cellular differentiation [24], investigating the immune checkpoint blockade [11] and so on [7–9]. These results show that DNBs can not only quantify early-warning signals before the disease onset or disease deterioration, but also discover dynamic biomarkers that can predict prognosis of each patient, thus having great potential to achieve the real preventive medicine and the ultra-early precision medicine.

On the other hand, entropy, a fundamental concept and measure in the thermodynamic theory, also plays an important role in characterizing fluctuation and correlation among the variables of a system. Recently, based on the generic DNB properties, a series of network-entropy-based computational methods have been proposed, to quantitatively detect the collective fluctuation of a local network before the critical transition [15,16]. In contrast to traditional methods which diagnose a disease state on the bases of differential expressions of biomolecules, these methods explore the information of differential associations or differential networks among biomolecules, and thereby are capable to predict a disease state or diagnose a pre-disease state. However, the original DNB method essentially requires data of multiple samples for each individual, which are unavailable for many cases. To further apply the DNB method to clinic data or for clinic practice, single-sample based methods were also developed to detect the pre-disease states on a single sample basis by measuring differential probability distributions or Kullback–Leibler divergence [19,22,30]. Specifically, by characterizing the perturbations of both correlation between molecules and variation of each molecule brought by each individual sample against a group of given control/reference samples, these methods exploit the differential distribution of observed high-dimensional biomolecules, which reflects the collective perturbations on the network. With such information derived from each individual, local network entropy (LNE) was proposed to construct the network entropy landscape, which was successfully applied for monitoring the outbreaks of COVID-19 and other infectious diseases [3,20]. Those works also indicate that we expect to detect probability distribution transitions rather than state transitions when strong noises or fluctuations are present in the observed data [14]. Clearly, this entropy-based approach is related to thermodynamic models of adaption shown in [6]. Actually, in addition to local bifurcation or the parameter tipping, rate tipping, noise tipping and distribution tipping are also observed and are studied in both theoretical and numerical perspectives [14].

Different from homogeneous systems, collective fluctuation for a heterogeneous biological system generally starts locally or at a local subnetwork, and then results in the state/distribution transition globally. Such collective fluctuation can be represented in the form of a network, i.e. correlation and variation, or second-order moments. Actually, more and more successful applications suggest that the dynamics-based and network-based methods already become essential techniques that allow us to probe into the underlying mechanism of complex biological systems [6], and may also be promising in developing effective clinical early-warning system. Moreover, among a number of intertwined variables/biomolecules among a dynamical network or DNB, revealing their causal relations [10] is also of great importance in elucidating molecular mechanisms of disease progression and designing appropriate drug therapy, which demands both statistics-based data science and dynamics-based data science [1,4,26] including deep learning and artificial intelligence.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the National Key R&D Program of China (No. 2017YFA0505500), the Strategic Priority Research Program of the Chinese Academy of Sciences (No. XDB38040400), National Natural Science Foundation of China (Nos. 11771152, 12026608 and 31930022), Guangdong Basic and Applied Basic Research Foundation (No. 2019B151502062), Japan Society for the Promotion of Science KAKENHI (No. 15H05707), Japan

Science and Technology Agency Moonshot R&D (No. JPMJMS2021), Japan Agency for Medical Research and Development (No. JP20dm0307009), UTokyo Center for Integrative Science of Human Behavior (CiSHuB), and by the International Research Center for Neurointelligence (WPI-IRCN) at The University of Tokyo Institutes for Advanced Study (UTIAS).

References

- [1] Chen C, Li R, Shu L, et al. Predicting future dynamics from short-term time series using an Anticipated Learning Machine. *Nat Sci Rev* 2020;7:1079–91.
- [2] Chen L, Liu R, Liu ZP, Li M, Aihara K. Detecting early-warning signals for sudden deterioration of complex diseases by dynamical network biomarkers. *Sci Rep* 2012;2:1–8.
- [3] Chen P, Chen E, Chen L, Zhou XJ, Liu R. Detecting early-warning signals of influenza outbreak based on dynamic network marker. *J Cell Mol Med* 2019;23:395–404.
- [4] Chen P, Liu R, Aihara K, Chen L. Autoreservoir computing for multistep ahead prediction based on the spatiotemporal information transformation. *Nat Commun* 2020;11:1–15.
- [5] Chen P, Liu R, Li Y, Chen L. Detecting critical state before phase transition of complex biological systems by hidden Markov model. *Bioinformatics* 2016;32:2143–50.
- [6] Gorban AN, Tyukina TA, Pokidysheva LI, et al. Dynamic and thermodynamic models of adaptation. *Phys Life Rev* 2021;37:17–64 [in this issue].
- [7] Jiang Z, Lu L, Liu Y, Zhang S, Li S, Wang G, et al. SMAD7 and SERPINE1 as novel dynamic network biomarkers detect and regulate the tipping point of TGF-beta induced EMT. *Sci Bull* 2020;65:842–53.
- [8] Koizumi K, Oku M, Hayashi S, Inujima A, Shibahara N, Chen L, et al. Identifying pre-disease signals before metabolic syndrome in mice by dynamical network biomarkers. *Sci Rep* 2019;9:1–11.
- [9] Koizumi K, Oku M, Hayashi S, Inujima A, Shibahara N, Chen L, et al. Suppression of dynamical network biomarker signals at the predisease state (Mibyō) before metabolic syndrome in mice by a traditional Japanese medicine (Kampo formula) Bofutsushosa. *Evid-based Compl Alt* 2020;2020:9129134.
- [10] Leng S, Ma H, Kurths J, Lai YC, Lin W, Aihara K, et al. Partial cross mapping eliminates indirect causal influences. *Nat Commun* 2020;11(1):1–9.
- [11] Lesterhuis WJ, Bosco A, Millward MJ, et al. Dynamic versus static biomarkers in cancer immune checkpoint blockade: unravelling complexity. *Nat Rev Drug Discov* 2017;16:264–72.
- [12] Li M, Liu Z, Qian B, Liu W, Horimoto K, Zhou H, et al. “Dysfunctions” induced by Roux-en-Y gastric bypass surgery are concomitant with metabolic improvement independent of weight loss. *Cell Discov* 2020;6:1–17.
- [13] Li M, Zeng T, Liu R, Chen L. Detecting tissue-specific early warning signals for complex diseases based on dynamical network biomarkers: study of type 2 diabetes by cross-tissue analysis. *Brief Bioinform* 2014;15:229–43.
- [14] Liu R, Chen P, Aihara K, Chen L. Identifying early-warning signals of critical transitions with strong noise by dynamical network markers. *Sci Rep* 2015;5:1–13.
- [15] Liu R, Chen P, Chen L. Single-sample landscape entropy reveals the imminent phase transition during disease progression. *Bioinformatics* 2020;36:1522–32.
- [16] Liu R, Li M, Liu Z-P, Wu J, Chen L, Aihara K. Identifying critical transitions and their leading biomolecular networks in complex diseases. *Sci Rep* 2012;2:813.
- [17] Liu R, Wang J, Ukai M, Sewon K, Chen P, Suzuki Y, et al. Hunt for the tipping point during endocrine resistance process in breast cancer by dynamic network biomarkers. *Cell Mol Biol* 2019;11:649–64.
- [18] Liu R, Wang X, Aihara K, Chen L. Early diagnosis of complex diseases by molecular biomarkers, network biomarkers, and dynamical network biomarkers. *Med Res Rev* 2014;34:455–78.
- [19] Liu R, Yu X, Liu X, Xu D, Aihara K, Chen L. Identifying critical transitions of complex diseases based on a single sample. *Bioinformatics* 2014;30:1579–86.
- [20] Liu R, Zhong J, Hong R, Chen E, Aihara K, Chen P, et al. Predicting local COVID-19 outbreaks and infectious disease epidemics based on landscape network entropy. *Sci Bull* 2021.
- [21] Liu X, Chang X, Leng S, Tang H, Aihara K, Chen L. Detection for disease tipping points by landscape dynamic network biomarkers. *Nat Sci Rev* 2019;6:775–85.
- [22] Liu X, Chang X, Liu R, Yu X, Chen L, Aihara K. Quantifying critical states of complex diseases using single-sample dynamic network biomarkers. *PLoS Comput Biol* 2017;2017:13:e1005633.
- [23] Mojtahedi M, Skupin A, Zhou J, Castaño IG, Leong-Quong RY, Chang H, et al. Cell fate decision as high-dimensional critical state transition. *PLoS Biol* 2016;14:e2000640.
- [24] Richard A, Boullu L, Herbach U, Bonnafoux A, Morin V, Vallin E, et al. Single-cell-based analysis highlights a surge in cell-to-cell molecular variability preceding irreversible commitment in a differentiation process. *PLoS Biol* 2016;14:e1002585.
- [25] Schadt EE. Molecular networks as sensors and drivers of common human diseases. *Nature* 2009;461:218–23.
- [26] Shi J, Aihara K, Chen L. Dynamics-based data science in biology. *Nat Sci Rev* 2021.
- [27] Tang H, Tang Y, Zeng T, Chen L. Gene expression analysis reveals the tipping points during infant brain development for human and chimpanzee. *BMC Genomics* 2020;21:1–10.
- [28] Van Nes EH, Scheffer M. Slow recovery from perturbations as a generic indicator of a nearby catastrophic shift. *Am Nat* 2007;169:738–47.

- [29] Yang B, Li M, Tang W, Liu W, Zhang S, Chen L, et al. Dynamic network biomarker indicates pulmonary metastasis at the tipping point of hepatocellular carcinoma. *Nat Commun* 2018;9:1–14.
- [30] Zhong J, Liu R, Chen P. Identifying critical state of complex diseases by single-sample Kullback–Leibler divergence. *BMC Genomics* 2020;21:87.

Comment

Some aspects of adaptation and evolution

Comment on “Dynamic and thermodynamic models of adaptation”

by A.N. Gorban et al.

V.G. Red’ko

Scientific Research Institute for System Analysis, Russian Academy of Sciences, Nakhimovsky Prospect, 36, build. 1, 117218, Moscow, Russia

Received 13 April 2021; accepted 13 April 2021

Available online 15 April 2021

Communicated by L. Perlovsky

The article by Gorban et al. [1] characterizes different aspects of adaptation. In particular, physical methods of analysis of complex systems are used to characterize adaptation.

The introduction describes the general approaches of the article. Section 2 is devoted to the analysis of *correlations* and *variation* of the relevant data from different areas. Medical, physiological, biological, climate, and human social systems are analyzed by means of considered methods. *Correlation graphs* are used in this analysis. Special attention is paid to problems of crises and risks. In particular, methods of obtaining *early warning indicators* of stress and crisis are described. This is essentially important for medical applications: for using at myocardial infarction, cancer, chronic hepatitis. Methods of revealing ‘*no return*’ points before patient death are characterized. The analysis of the article [1] sheds light on these methods and important medical problems. Ecological and human social systems, as well as, markets and financial systems are also analyzed by means of correlation graphs. The analysis of Section 2 really demonstrates that “construction of early warning indicators of stress, illness and various crisis on the basis of correlation graphs and variance of relevant attributes is possible [1].”

Section 3 of the article [1] develops previous works by Selye [2,3], Cannon [4], and Goldstone [5,6]. Namely, certain thermodynamics of adaptation is proposed and analyzed in this section. The notions “*adaptation energy*”, “*adaptation resource*” are used in this analysis. Methods of choosing the most important factors and optimization of resource allocation are especially analyzed. It is interesting that the authors of the reviewed article introduced the notion “*instantaneous fitness*” for “... measuring an individual performance in a highly non-trivial task”. It is underlined that phenotype plays an essential role in the selection of organisms into future generations. Phenotypes can be adjusted during the adaptation of organisms in life. Certain approaches to investigations of phenotype role in evolutionary processes were developed by Hinton and Nowlan, Maley, and Red’ko in works devoted to studying the interaction between learning and evolution [7–10]. These works will be outlined briefly below.

The dynamics of individual adaption are characterized in Section 4. The model of effective use of resources in processes of adaption is developed in this section.

DOI of original article: <https://doi.org/10.1016/j.plrev.2021.03.001>.

E-mail address: vgredko@gmail.com.

Section 5 concludes the article by Gorban et al.

The reviewed article is interesting and important. However, it has certain small imperfections. In particular, Section 2 contains good summarizing notes, but Sections 3, 4 have not such summary that would be useful for readers of the article.

Interesting directions of possible investigations that are related to the reviewed article are outlined below.

The authors of the reviewed article consider several particular problems of the development of social systems. This analysis is important, however, there is a possibility to analyze the more global problems of development of human society. Currently, there are numerous international and interstate conflicts in the world. These conflicts lead to local wars, armed clashes, and interstate tension. How could we overcome the tendencies towards interethnic confrontation, to international and interstate conflicts? According to the author of this comment, the most natural way to reduce international tension is the all-around development of mutual understanding between different countries and peoples.

A certain example of a harmonious united society with different structures and with different languages is Switzerland, a federal republic consisting of 20 cantons. Each canton has its own constitution and legislation, but their effect is limited by the federal constitution. It is possible that this example of the successful unification of different groups of people, with different habits and lifestyles, will be useful in building a global harmonious humankind community.

Concerning the instantaneous fitness analyzed in [1], let us characterize the relation between phenotype and genotype at the interaction between learning and evolution. Such interaction was investigated in works [7–10]. The evolving population of agents was considered. Any agent has the genotype and phenotype. At the beginning of generation, the phenotype of an agent is determined by its genotype. During a generation, the genotypes of agents are not changed, whereas the agent phenotypes are optimized by means of learning. At the generation ends, the fitness of each agent is determined. The selection of agents for the next population is based on their fitness, which is determined by the final phenotypes they acquire via the learning during the generation. The genotypes of selected agents are transmitted (with small mutations) from the agent-parents to the agent-children at forming the next generation. As descendants of organisms obtain genotypes that organisms received from their parents and not phenotypes, the evolutionary process has Darwinian character.

The works [7–10] demonstrated the following effects of the interaction between learning and evolution:

- 1) The genetic assimilation of abilities acquired by individual learning (during a large number of generations).
- 2) The hiding effect by which intensive learning can retard the evolutionary optimization of genotypes.
- 3) The effect of learning load (i.e. the impairment of agent fitness caused by a considerable change of the phenotype in learning), which can quicken the evolutionary optimization.

The agent phenotypes obtained via learning in [7–10] are similar to instantaneous fitness (obtained via adaptation) considered in [1].

Declaration of competing interest

The author declares that he has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work was supported by the State Program of SRISA RAS. Project No. 0065-2019-0003.

References

- [1] Gorban AN, Tyukina TA, Pokidysheva LI, Smirnova EV. Dynamic and thermodynamic models of adaptation. *Phys Life Rev* 2021;37:17–64 [in this issue].
- [2] Selye H. Adaptation energy. *Nature* 1938;141(3577):926.
- [3] Selye H. Experimental evidence supporting the conception of “adaptation energy”. *Am J Physiol* 1938;123:758–65.
- [4] Cannon WB. *The wisdom of the body*. NY: W.W. Norton; 1932.
- [5] Goldstone B. The general practitioner and the general adaptation syndrome. *South Afr Med J* 1952;26:88–92. PMID: 14901129.
- [6] Goldstone B. The general practitioner and the general adaptation syndrome. Concluded. *South Afr Med J* 1952;26:106–9. PMID: 14913266.
- [7] Hinton GE, Nowlan SJ. How learning can guide evolution. *Complex Syst* 1987;1(3):495–502.

- [8] Mayley G. Guiding or hiding: explorations into the effects of learning on the rate of evolution. In: Husbands P, Harvey I, editors. *Proceedings of the fourth European conference on artificial life*. Cambridge, Massachusetts: The MIT Press; 1997. p. 135–44.
- [9] Red'ko VG. Mechanisms of interaction between learning and evolution. *Biol Inspir Cognit Archit* 2017;22:95–103.
- [10] Red'ko VG. Spin glass energy minimization through learning and evolution. *Opt Mem Neural Netw (Inf Opt)* 2020;29(3):187–97.

Comment

Variational free energy, individual fitness, and population dynamics
under acute stressComment on “Dynamic and thermodynamic models of adaptation”
by Alexander N. Gorban et al.Kai Ueltzhöffer^{a,b,*}, Lancelot Da Costa^{a,c}, Karl J. Friston^a^a Wellcome Centre for Human Neuroimaging, University College London, UCL Queen Square Institute of Neurology, 12 Queen Square, London WC1N 3AR, United Kingdom^b Department of General Psychiatry, Centre of Psychosocial Medicine, Heidelberg University Hospital, Voßstraße 2, 69115 Heidelberg, Germany^c Department of Mathematics, Imperial College London, London SW7 2AZ, United Kingdom

Received 15 April 2021; accepted 15 April 2021

Available online 19 April 2021

Communicated by L. Perlovsky

When we agreed to comment on the intriguing review by Gorban et al. [1], we initially assumed it would address the statistical physics of dissipative adaptation, which aims to develop an account of self-organisation of active matter by building on recent advances in non-equilibrium statistical physics [2]. We were pleasantly surprised that the target article beautifully summarized data and theories from—at first glance—a very different field: namely, the adaptation of populations of complex systems (e.g., biological individuals) under acute stress. However, on deeper inspection, there are many interesting points of contact between these fields. For example, Horowitz et al. [3] calculate the minimum work rate required (i.e., the minimum amount of thermodynamic free energy that a system has to consume per unit time) to maintain a specific non-equilibrium (“living”) steady state, which directly connects to Appendix A of the target article.

This commentary will focus on another body of work, which speaks to the treatment of Gorban et al. [1], namely, the free energy principle [4,5]. The free-energy principle stipulates that adaptive systems (characterised as certain non-equilibrium steady-states [6]) embody a generative model of their environment that entails ‘preferred’ (i.e., characteristic, system-defining) sensations. Furthermore, it assumes that these systems perform ‘perception’, ‘learning’ and ‘action’ by minimising a quantity known as variational free energy. Variational free energy is a functional optimised in variational Bayesian inference—known as the negative evidence lower bound (ELBO) in machine learning and statistics [7]. A formal definition is given in Box 1. In short, it is a tractable upper-bound on an agent’s surprise under their generative model of the world. This upper-bound is minimised via perception and learning, which correspond to variational Bayesian inference on the state and contingencies of the world, respectively [8–10]. Complementarily, this bound is minimised via action, which controls the agent towards preferred sensations [11,12]. As the

DOI of original article: <https://doi.org/10.1016/j.plrev.2021.03.001>.

* Corresponding author at: Wellcome Centre for Human Neuroimaging, University College London, UCL Queen Square Institute of Neurology, 12 Queen Square, London WC1N 3AR, United Kingdom.

E-mail address: kai.ueltzhoeffer@med.uni-heidelberg.de (K. Ueltzhöffer).

agent's optimism-biased generative model favours 'familiar' homeostatic states—by assigning them a high a-priori probability—the agent actively seeks out these homeostatic states, responding adaptively to (potentially stressful) perturbations from the environment. This adaptive mechanism might be implemented via simple homeostatic reflexes, higher-order allostatic processes, and even sophisticated counterfactual planning to prevent or escape potentially life-threatening situations with long-term consequences [13,14]. This kind of 'active' Bayesian inference is called 'active inference'. Given a hierarchical, nonlinear, and dynamic generative model, and a simple choice of variational family and inference procedure, the variational free energy reduces to a sum of precision-weighted prediction errors, reproducing central features of hierarchical predictive coding formulations of brain architecture and function [14–17].

A central question raised by Gorban et al. [1], is how one could quantify the instantaneous individual fitness (or wellbeing W) of an agent. One might think that the variational free energy is in a good position to score wellbeing. It approximates an agent's instantaneous surprisal about sensations—under their optimism-biased generative model of the world—featuring prior expectations of finding themselves in a homeostatic state, without pain, freezing, hunger, thirst, dyspnoea, etc. Thus, a low free energy corresponds to an instantaneous state of the agent, which is well aligned with their prior expectations to find themselves within their 'comfort zone'. Thereby, their variational free energy directly quantifies (the inverse of) their instantaneous wellbeing by measuring the agent's instantaneous 'distress'.

Furthermore, the dynamics of active inference, under the free energy principle, might also explain the core of empirical findings summarized by Gorban et al. [1], namely, the simultaneous increase in the variance and correlations of physiological parameters in populations under acute stress due to an elongation of the distribution in state space along a single dimension (corresponding to a "bullet-shape", as shown in Figure 2 in [1]). In Fig. 1, we show that a population of simple active inference agents—two dimensional "Bayesian thermostats"—exhibit this canonical pattern of transiently increasing correlation and variance under acute stress. The simulated agents were displaced from one environment, to which they were well adapted, into another environment: similar to the children born in the temperate zone of Siberia versus the far north in [1]. This displaced the two homeostatic variables X and Y of the agents (e.g., body temperature and blood pH) from their comfortable range, close to the origin, to a distal value. Here, we assume that this displacement is in the same direction in state space for all agents: i.e., all agents are subject to similar stressful conditions. To minimise their variational free energy, the agent's actions will compensate these changes, to bring back X and Y to their comfort zone around the origin (by minimising an appropriate interoceptive prediction error). Crucially, we assume that the conspecifics all feature a generative model from the same model family (e.g., species). Thus, the most likely trajectories of individuals back to the comfort zone follow a narrow channel of least (variational) action [14]. Furthermore, in our simulations, we introduced individual variability in the agents' generative models, corresponding to slight differences in the time-scales of their internal compensatory mechanisms. This can be thought of, for example, as slower or faster metabolisms, or stronger or weaker effector organs. Thereby, the agents start to disperse (due to individual variations) during their transition via a narrow channel (due to their shared model structure) leading them back to their homeostatic comfort zone. These dynamics directly reproduce the "bullet shape" distribution shown in Figure 2 of [1] and accord nicely with the compelling evidence summarized in [1].

A final, but important point to comment on, is the relationship between variational free energy and thermodynamic free energy. Because the variational free energy approximates the surprisal under an optimism-biased generative model of the world, it measures instantaneous distress, or the misalignment of current outcomes with preferred outcomes. Crucially, this means that the minimisation of variational free energy in living agents is associated with the build-up and maintenance of the physical structure realizing these agents. This macroscopic structure exists far from thermodynamic equilibrium. Accordingly, the corresponding macroscopic disequilibria actually store thermodynamic free energy [18], leading to an inverse relationship between the two quantities. Just think of a small seed, which turns into a large tree during its development and growth. On the one hand, this system can be described as trying to fulfil the expectations encoded epigenetically in its embodied generative model of the world, via minimizing the corresponding variational free energy [19]. On the other hand, the seed turns into a formidable, large reservoir of thermodynamic free energy, which could be harvested to heat a home during winter (for details see pp. 118–121 of [20]).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Box 1: Variational free energy and active inference

The variational free energy F depends on a generative model $p(o, x, \theta) = p(o | x, \theta)p(x | \theta)p(\theta)$ of observations o , latent states of the environment x and parameters θ (including hyperpriors), describing the dynamic structure of the environment. Furthermore, the variational free energy depends on a variational distribution $q(x, \theta | \mu)$ on the latent states and parameters, which is usually chosen from a parameterised family of distributions. This allows it to be specified by a small set of sufficient statistics μ (e.g., the mean and covariance matrix of normal distributions).

The variational free energy can now be defined as:

$$\begin{aligned}
 F(\mu, o) &= \underbrace{-\ln p(o)}_{\text{Surprisal}} + \text{D}_{\text{KL}}(q(x, \theta | \mu) || p(x, \theta | o)) \\
 &= \underbrace{-(\ln p(o | x, \theta))_{q(x, \theta | \mu)}}_{\text{Accuracy}} + \underbrace{\text{D}_{\text{KL}}(q(x, \theta | \mu) || p(x, \theta))}_{\text{Complexity}} \\
 &= \langle -\ln p(o, x, \theta) \rangle_{q(x, \theta | \mu)} - H(q(x, \theta | \mu))
 \end{aligned}$$

$\text{D}_{\text{KL}}(q(x) || p(x)) = \int_x q(x) \ln \frac{q(x)}{p(x)} dx$ denotes the Kullback-Leibler divergence between two distributions $p(x)$ and $q(x)$, which is always non-negative, and zero if and only if p and q are identical distributions [21,22]. Thus, the first line shows that the variational free energy is indeed an upper bound on the surprisal, i.e., the negative log evidence of the observations o under the generative model $p(o, x, \theta)$. Furthermore, it shows that minimising the variational free energy brings the variational density closer to the true posterior over latent states and parameters, given sensations, thereby licensing the interpretation as approximate Bayesian inference. However, the calculation of the surprisal term in the first line is intractable in general as it involves the marginalisation $p(o) = \int_{x, \theta} p(o, x, \theta) dx d\theta$ over the (potentially extremely high-dimensional) latent spaces x and θ . By using the basic properties of integrals and the logarithm, it is straightforward to show that the variational free energy can also be cast in terms of the negative of the expected log likelihood under the variational distribution, called “accuracy”, plus the Kullback-Leibler divergence between the variational density and the prior on latent variables, called “complexity”, as shown in the second line of the above equation. While the interpretation of accuracy—namely, how likely the data is under the variational posterior—is straightforward, the complexity term is more nuanced: complexity measures the discrepancy between prior and posterior beliefs. This means that a good explanation, in terms of a variational posterior, provides an accurate account of empirical data (maximising accuracy), which requires minimal changes during the updating of prior to posterior beliefs (minimising complexity, cf., Occam’s principle). Furthermore, by Landauer’s principle a change in beliefs entails heat generation, therefore complexity also bounds the minimum thermodynamic cost of inference [23]. From a practical standpoint, given a suitable choice of the variational density, e.g., in terms of normal distributions, these expressions can be easily evaluated, thus converting a very complex integration problem into a much easier optimisation problem. The last line is another rearrangement of the variational free energy, showing a formal similarity with thermodynamic free energy. Please bear in mind that this similarity only exists in the functional form, since in reality, variational free energy and thermodynamic free energy of self-organising, living systems often display a reciprocal relationship. Variational free energy minimisation via the sufficient statistics describing the distribution over latent states x corresponds to perception, while minimisation via the sufficient statistics describing the distribution over parameters θ corresponds to approximate Bayesian inference on environmental contingencies, i.e., learning [9,24]. Crucially, living agents influence their environment through actions a , which control the observations they sample $o(a)$. Thereby, such “active inference” agents can actively sample observations with high likelihood under their optimism-biased generative model, thereby ensuring their continued existence via self-evidencing [14]. To link back to [1], active inference agents under acute stress feature a high variational free energy, which in turn implies either unexpected, unpleasant sensations (low accuracy) or a high Kullback-Leibler divergence between their variational posterior estimate of their current state and their a priori (“preferred”) state (high complexity), or both.

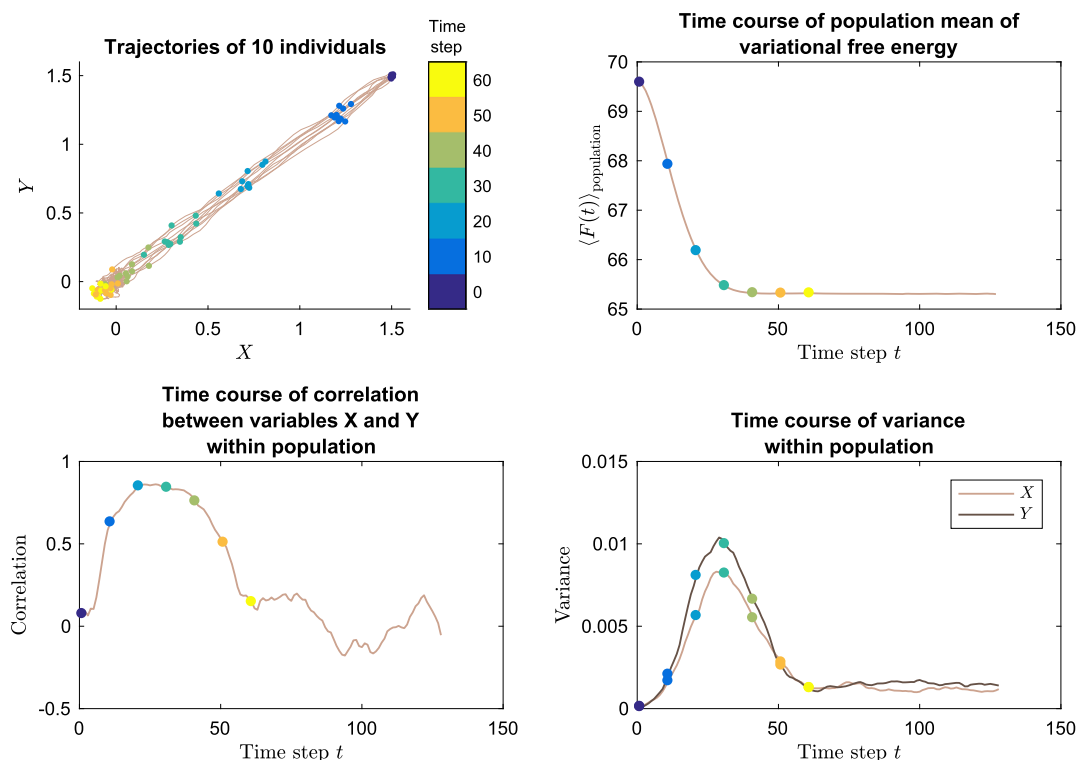


Fig. 1. Transient increase in the correlation and variance within an ensemble of active inference agents under acute stress. We simulated a population of 100 active inference agents, featuring two latent state (“physiological”) variables X and Y , conforming to a continuous-time, continuous-state-space version of active inference [14]. The agents are assumed to have suffered from an abrupt displacement from their “comfortable” homeostatic state (close to the origin $X = Y = 0$) at the initial timestep. Following this perturbation, homeostatic reflex arcs elicit actions to minimize interoceptive prediction errors, restoring the physiological variables to their expected, homeostatic range. Due to individual variability, the time-constants of their internal dynamics are slightly different, leading to a spread of the population across the (one-dimensional) channel corresponding to the path of least (variational) action back to their “comfortable zone”. The upper left panel plots ten exemplary trajectories (beige lines), with the state of the ten agents plotted for seven different time points (coloured dots). This shows how the ensemble transitions from an initially isotropic shape via an elliptical (“bullet-shaped”) distribution back to an isotropic shape. The upper right panel shows the temporal evolution of the mean variational free energy of the population, i.e., the mean of individual agents’ current “distress”, as their actions move them back to their “comfort zone” around the origin. The corresponding evolution of the correlation between the two physiological state variables, and their individual variances are shown in the lower panels. The simulation code producing this figure is available via https://github.com/kaiu85/PoLR_comment, it is based on the code accompanying [14], which is freely available as part of the Statistical Parametric Mapping (SPM) software (<https://www.fil.ion.ucl.ac.uk/spm/software/>).

Acknowledgements

Kai Ueltzhöffer is supported by a DAAD PRIME fellowship. Lancelot Da Costa is supported by the Fonds National de la Recherche Luxembourg (Project code: 13568875). Karl J. Friston is a Wellcome Principal Research Fellow (Ref: 088130/Z/09/Z).

References

- [1] Gorban AN, et al. Dynamic and thermodynamic models of adaptation. *Phys Life Rev* 2021;37:17–64 [in this issue].
- [2] Perunov N, Marsland RA, England JL. Statistical physics of adaptation. *Phys Rev X* 2016;6.
- [3] Horowitz JM, Zhou K, England JL. Minimum energetic cost to maintain a target nonequilibrium state. *Phys Rev E* 2017;95(4):042102.
- [4] Friston K, Kilner J, Harrison L. A free energy principle for the brain. *J Physiol (Paris)* 2006;100(1):70–87.
- [5] Friston K. The free-energy principle: a unified brain theory? *Nat Rev Neurosci* 2010;11(2):127–38.
- [6] Parr T, Da Costa L, Friston K. Markov blankets, information geometry and stochastic thermodynamics. *Philos Trans R Soc A, Math Phys Eng Sci* 2020;378(2164):20190159.
- [7] Blei DM, Kucukelbir A, McAuliffe JD. Variational inference: a review for statisticians. *J Am Stat Assoc* 2017;112(518):859–77.

- [8] Whittington JCR, Bogacz R. An approximation of the error backpropagation algorithm in a predictive coding network with local Hebbian synaptic plasticity. *Neural Comput* 2017;29(5):1229–62.
- [9] Buckley CL, et al. The free energy principle for action and perception: a mathematical review. *J Math Psychol* 2017;81:55–79.
- [10] Da Costa L, et al. Active inference on discrete state-spaces: a synthesis. *J Math Psychol* 2020;99:102447.
- [11] Baltieri M, Buckley CL. PID control as a process of active inference with linear generative models. *Entropy* 2019;21(3).
- [12] Ueltzhoffer K. Deep active inference. *Biol Cybern* 2018;112(6):547–73.
- [13] Corcoran AW, Pezzulo G, Hohwy J. From allostatic agents to counterfactual cognisers: active inference, biological regulation, and the origins of cognition. *Biol Philos* 2020;35(3):32.
- [14] Friston KJ, et al. Action and behavior: a free-energy formulation. *Biol Cybern* 2010;102(3):227–60.
- [15] Friston K, Kiebel S. Predictive coding under the free-energy principle. *Philos Trans R Soc Lond B, Biol Sci* 2009;364(1521):1211–21.
- [16] Friston K. Hierarchical models in the brain. *PLoS Comput Biol* 2008;4(11):e1000211.
- [17] Friston K. A theory of cortical responses. *Philos Trans R Soc Lond B, Biol Sci* 2005;360(1456):815–36.
- [18] Branscomb E, Russell MJ. Turnstiles and bifurcators: the disequilibrium converting engines that put metabolism on the road. *Biochim Biophys Acta, Bioenerg* 2013;1827(2):62–78.
- [19] Friston K, et al. Knowing one's place: a free-energy approach to pattern regulation. *J R Soc Interface* 2015;12(105):20141383.
- [20] Friston K. A free energy principle for a particular physics. *arXiv:1906.10184*, 2019.
- [21] Kullback S, Leibler RA. On information and sufficiency. *Ann Math Stat* 1951;22(1):79–86.
- [22] Cover TM, Thomas JA. *Elements of information theory*. New York: John Wiley & Sons; 2006.
- [23] Landauer R. Irreversibility and heat generation in the computing process. *IBM J Res Dev* 1961;5(3):183–91.
- [24] Bogacz R. A tutorial on the free-energy framework for modelling perception and learning. *J Math Psychol* 2017;76:198–211.

Comment

Aging, Inflammaging and Adaptation

Comment on “Dynamic and thermodynamic models of adaptation” by A.N. Gorban et al.

Claudio Franceschi

Department of Applied Mathematics, Mathematics of Future Technologies Center, Laboratory of Systems Medicine of Healthy Aging,
Lobachevsky University, Nizhny Novgorod 603950, Russia

Received 8 July 2021; accepted 8 July 2021

Available online 15 July 2021

Communicated by L. Perlovsky

Gorban et al. [1] present an up-to-date and insightful network approach to several biological and pathological conditions, including aging, among others. Even if my knowledge does not allow me to enter in technical details and evaluate the mathematical approach, as Editor-in-Chief of Aging Research Reviews, a top journal in the field of aging and age-related diseases, I like to express my gratitude to Gorban and coworkers for their conceptual effort and suggestions. Moreover, I predict that in the next future a dialogue/collaboration and deep intertwining between physicians/biologists and physicists/mathematicians like Gorban and coworkers will represent one the most promising development and perspective for the medicine of the elderly. The approach of this paper will help not only in disentangling the complexity of aging (particularly in humans) but also in establishing a new “aging medicine”, and above all in identifying new tools capable of early predicting power regarding human health trajectories. The paper of Gorban et al., is particularly interesting for me as an MD and as an immunologist who, in my long scientific activity, studied contemporarily both aging (with particular attention to healthy aging and centenarians), and many relevant age-related diseases (ARDs) and geriatric syndromes (GSs) (such as Alzheimer’s dementia, Parkinson’s disease, diabetes, cancer, frailty/sarcopenia, metabolic syndrome, among others).

Moreover, I greatly appreciated one of the conclusions of Gorban’s paper, when applications of the adaptation theory developed in the paper are suggested, and it is explicitly said that “*successful aging is not connected with excessive capacity to cope with stressors but rather those with capability to afford intermediate response to stress, i.e. neither too weak nor too strong, in relation to the environment. In terms of AE (Adaptation Energy, a key concept of the Gorban’s paper) distribution, this moderation can be seen as optimal use of AE for the purpose of longevity*”.

This conclusion fits very well with what I learned from the study of human aging and longevity in centenarians (100+), semi-supercentenarians (105+) and supercentenarians (110+) [2].

The main reasons of my deep interest can be summarized according to the following major steps and considerations which stem from one of the conclusion/recommendations of the Gorban’s et al. paper: “*special attention should be paid to such universal physiological factors as inflammation, which can be considered as a universal mechanism for accelerating aging (inflammaging)*”:

DOI of original article: <https://doi.org/10.1016/j.plrev.2021.03.001>.

E-mail address: claudio.franceschi@unibo.it.

1. Since the beginning of my studies on aging, and with all the limits of my medical and biological expertise, I was attracted by a global/systemic/mathematical approach. Indeed, I initially proposed two general theories, i.e., “the network” (1989, 2000) [3,4] and “the remodeling” theories of aging (1996, 2000) [4,5], which later/finally converged on the “INFLAMMAGING” theory of aging.

2. The NETWORK THEORY of aging suggests that in the course of evolution a number of mechanisms [DNA repair, production of heat shock and other stress proteins, enzymatic and non-enzymatic antioxidant defense systems, poly ADP-ribose polymerase activation, and cell death/apoptosis] have emerged which allow the cell to cope with a variety of potentially harmful agents and on the whole can be considered the basic molecular/cellular anti-aging system, being interconnected and constituting a NETWORK of cellular defense systems.

3. The REMODELING THEORY of aging suggests that: 1) aging and longevity are related in a complex way to the capability of the body to cope with and adapt to a variety of stressors (see point 7, below); 2) IMMUNOSENESCENCE (i.e., the aging of the immune system) is the net result of the continuous adaptation of the body to the deteriorative changes occurring over time. Accordingly, body resources are continuously optimized, and immunosenescence has to be considered a very dynamic process (including both loss and gain) deeply related to and interconnected with the age-related remodeling of metabolic pathways and endocrine functions.

4. The INFLAMMAGING theory starts from the consideration that aging is the major risk factor for ARDs and GSs, both sharing a common set of few basic mechanistic “pillars” (adaptation to stress, epigenetics, inflammation, macromolecular damage, metabolism, proteostasis, stem cells and regeneration) including inflammation [6]. It is crucial to note that with advancing age, and particularly in the post-reproductive period of life, mostly non predicted by evolution, the deterioration (molecular and functional increase and/or decrease) of these pillars converge on inflammation, which becomes a major driving force of the aging process. Inflammation is a basic, *beneficial* phenomenon crucial for survival, fueled by a multitude of stimuli which occur lifelong and are sensed by a limited set of evolutionary-selected promiscuous sensors/receptors whose “degeneracy” allows them to recognize a large variety of signals and trigger inflammatory responses by activating the evolutionary basic and ancestral “innate” immune system but also the evolutionary more recent and sophisticated “adaptive” immune system [7]. Inflammatory stimuli include bacterial and viral pathogens and their products (non-self), cell debris generated by the billions of cells which die every day, and by misplaced/altered molecules (self), as well as nutrients and GUT MICROBIOTA products (quasi-self).

5. However, INFLAMMATORY STIMULI AND RESPONSES INCREASE WITH AGE, THUS POTENTIALLY SHIFTING INFLAMMATION FROM BENEFICIAL TO DETRIMENTAL, eventually leading to a chronic, sterile, low-grade inflammation, that I proposed to name INFLAMMAGING [8]. Inflammaging, together with hormesis (i.e. a process in which exposure to a low dose of a chemical agent or environmental factor that is damaging at higher doses induces an adaptive beneficial effect on the cell or organism) frame the aging process within a lifelong perspective of ADAPTATION to different types of stressors [9]. Indeed, stressors can elicit responses with beneficial or detrimental effect depending on the intensity/duration, according to the hormesis principle. Such adaptation perspective also suggests that the aging process is malleable. It is important to note that both inflammaging and hormesis have a strong evolutionary basis and are highly energy-demanding processes which require strategic re-allocation of energy and profound metabolic changes, and that calorie restriction, the most thoroughly intervention capable of increasing the lifespan in a variety of organisms, has been conceptualized as a chronic hormetic stress [9].

6. Concomitantly, a decline of many immune function/parameters collectively termed “IMMUNOSENESCENCE” occurs with age. Together, inflammaging and immunosenescence are suggested to stand at the origin of ARDs and GSs, collectively conceptualized as “accelerated aging” [10].

7. I also argued that the universe of inflammatory stimuli fueling inflammaging largely coincides with that of STRESSORS, as conceptualized by Hans Selye [11]. From invertebrates to vertebrates, a large set of experimental data suggest that the stress response is an integrated immune-neuroendocrine response crucial for the hierarchical redistribution of resources within the body, evolved as adaptation to dynamic environmental conditions and specific ecological demands. I also proposed the hypothesis that a BOW-TIE architecture might be suitable to describe the variety of immune-neuroendocrine inflammatory inputs that continuously target cells and organs while, at the same time, minimizing the cost of immune-neuroendocrine responses [12]. Bow-tie architectures are able to convert a variety of stimuli (fan in) into a wide range of fine-tuned responses (fan out) by passing through the integrating activity of a core (knot) constituted by a limited number of elements.

8. According to this integrated view, aging/longevity and ARDs/GSs become part of A CONTINUUM where the two extremes are represented by centenarians (100+) and supercentenarians (110+), who largely avoided or postponed

most ARDs/GSs and are characterized by *decelerated* aging, and patients who suffered one or more severe ARDs/GSs and show signs of *accelerated* aging [13]. In between these two extremes there is a continuum of intermediate trajectories. Such large phenotypic heterogeneity depends on an extremely complex, lifelong interaction between the genetic background of each individual and the historically and geographically-specific environmental and lifestyle factors, and variables that each person adopted and/or has been exposed to lifelong. The major variables include, sex/gender (aging is largely different in men and women), education, economic status, type of work, nutrition, among others. To this regard, I proposed that a key concept to understand such huge heterogeneity of possible health trajectories is “IMMUNOBIOGRAPHY”, i.e., the unique combination of type, dose, intensity, and temporal sequence of antigenic/inflammatory stimuli (related to the above-mentioned set of variables) that each subject experienced lifelong, starting in utero [14]. The innate and adaptive immune systems of each individual are capable of recording and adapting to all the immunological experiences encountered lifelong, owing to their memory and plasticity. Accordingly, I proposed/conceptualized the “liquid immune self”, suggesting the need of a “personalized” approach, challenging the tenet that inflammaging and immunosenescence are basically only *negative* phenomena. I also argued that from an evolutionary perspective inflammaging and immunosenescence can be better viewed as ADAPTIVE rather than solely detrimental (resulting from the above-mentioned lifelong individually-based remodeling), being intriguingly involved not only in ARDs and GSs but also in survival and longevity [15].

To my knowledge, this conceptual scenario represents the basic and most advanced biological framework necessary to fully and critically appreciate the effort of Gorban et al. to find a common and general mathematical approach to the various examples of diseases addressed and discussed in his paper.

Finally, what about possible future development of Gorban et al. approach?

Few considerations:

1. As I noticed several times, a basic word of the Gorban’s et al. paper is ADAPTATION (an extension and rigorous refinement of Hans Selye “adaptability”), which is assumed as the key to merge biology and physics (thermodynamics), and for the development of the new discipline of “adaptation engineering”. I think that it is not by chance that ADAPTATION is also the basic conceptual background of inflammaging, as I summarized in the previous paragraphs.

2. One conclusion of the Gorban’s et al. paper is the “*application of the models of individual adaptation (Sec 4) to training and sport physiology*”. I not only agree but suggest to expand the application of the proposed dynamic and thermodynamic models to other topics such as nutrition, and also to all other ARDs and GSs.

3. A major usually non explicit assumption of the approach adopted by Gorban et al. is that the elements of each specific cellular network are basically similar each other’s. This basic assumption is likely not true. Recent genetic, epigenetic and immunological data *at a single cell level* in different organs of the body suggest that *each cell has a specific signature and present subtle differences with respect to the other cells derived from the same source*. These observations add a further, still largely unexplored level of biological complexity to the organs and systems of the body, and pose a variety of physiological and pathological questions at present largely unanswered.

4. The above considerations are even more crucial if we accept the fundamental heterogeneity and uniqueness of humans at all biological level (genetic, epigenetic, metabolic, among others) which emerges whenever the results of a study are analyzed at the individual level. Regarding aging, such heterogeneity increases dramatically with age [14,16], a phenomenon evident also at the psychological level, as recognized about a century ago by Carl Gustav Jung with his *principium individuationis* [17]. From this point of view an enormous amount of work is expecting both physicists, biologists and MDs.

5. Assuming that such heterogeneity at single cell level in the different organs and systems of the body is pervasive, the new possible questions from an aging point of view are the following: how this heterogeneity impact on the global function of different organs during aging? Does single cell heterogeneity change (increase? decrease?) with age? Are these possible age-related changes similar or different in the different organs of the body? How and how much the intercellular and inter-organs communication is affected? Which is the impact of the above-mentioned variables (gender, lifestyle. . .) on single cell heterogeneity and its change with age?

In my opinion these and many other possible questions represent an urgent challenge for scientists like Gorban and coworkers and their approach, as outlined in their stimulating paper.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The author acknowledges support by the grant of the Ministry of Education and Science of the Russian Federation Agreement No. 074-02-2018-330.

References

- [1] Gorban A, Tyukina T, Pokidysheva L, Smirnova E. Dynamic and thermodynamic models of adaptation. *Phys Life Rev* 2021;37:17–64. <https://doi.org/10.1016/j.plrev.2021.03.001>.
- [2] Franceschi C, Ostan R, Santoro A. Nutrition and inflammation: are centenarians similar to individuals on calorie-restricted diets? *Annu Rev Nutr* 2018;38:329–56.
- [3] Franceschi C. Cell proliferation, cell death and aging. *Aging (Milano)* 1989;1(1):3–15.
- [4] Franceschi C, Monti D, Barbieri D, Salvioli S, Grassilli E, Capri M, et al. Successful immunosenescence and the remodelling of immune responses with ageing. *Nephrol Dial Transplant* 1996;11(Suppl 9):18–25.
- [5] Franceschi C, Valensin S, Bonafè M, Paolisso G, Yashin AI, Monti D, et al. The network and the remodeling theories of aging: historical background and new perspectives. *Exp Gerontol* 2000;35(6–7):879–96.
- [6] Kennedy BK, Berger SL, Brunet A, Campisi J, Cuervo AM, Epel ES, et al. Geroscience: linking aging to chronic disease. *Cell* 2014;159:709–13.
- [7] Franceschi C, Garagnani P, Parini P, Giuliani C, Santoro A. Inflammaging: a new immune-metabolic viewpoint for age-related diseases. *Nat Rev Endocrinol* 2018;14(10):576–90.
- [8] Franceschi C, Bonafè M, Valensin S, Olivieri F, De Luca M, Ottaviani E, et al. Inflamm-aging. An evolutionary perspective on immunosenescence. *Ann NY Acad Sci* 2000;908:244–54.
- [9] Santoro A, Martucci M, Conte M, Capri M, Franceschi C, Salvioli S. Inflammaging, hormesis and the rationale for anti-aging strategies. *Ageing Res Rev* 2020;64:101142.
- [10] Franceschi C, Santoro A, Capri M. The complex relationship between immunosenescence and inflammaging. *Semin Immunopathol* 2020;42(5):517–20. Special issue on the New Biomedical Perspectives.
- [11] Ottaviani E, Franceschi C. The neuroimmunology of stress from invertebrates to man. *Prog Neurobiol* 1996;48(4–5):421–40.
- [12] Tieri P, Grignolio A, Zaikin A, Mishto M, Remondini D, Castellani GC, et al. Network, degeneracy and bow tie. Integrating paradigms and architectures to grasp the complexity of the immune system. *Theor Biol Med Model* 2010;7:32.
- [13] Franceschi C, Garagnani P, Morsiani C, Conte M, Santoro A, Grignolio A, et al. The continuum of aging and age-related diseases: common mechanisms but different rates. *Front Med (Lausanne)* 2018;5:61.
- [14] Franceschi C, Salvioli S, Garagnani P, de Eguileor M, Monti D, Capri M. Immunobiography and the heterogeneity of immune responses in the elderly: a focus on inflammaging and trained immunity. *Front Immunol* 2017;8:982.
- [15] Fulop T, Larbi A, Dupuis G, Le Page A, Frost EH, Cohen AA, et al. Immunosenescence and inflamm-aging as two sides of the same coin: friends or foes? *Front Immunol* 2018;8:1960.
- [16] Franceschi C, Garagnani P, Olivieri F, Salvioli S, Giuliani C. The contextualized genetics of human longevity. *JACC Focus Seminar J Am Coll Cardiol* 2020;75(8):968–79.
- [17] Jung Carl Gustav. *Coscienza, Inconscio e Individuazione*; 1935.

Comment

Critical transition across the Waddington landscape as an interpretative model

Comment on “Dynamic and thermodynamic models of adaptation” by A.N. Gorban et al.

Mariano Bizzarri ^{*}, Paola Pontecorvi*Dept. of Experimental Medicine, Sapienza University of Rome, Systems Biology Group Lab, Rome, Italy*

Received 28 May 2021; accepted 31 May 2021

Available online 6 June 2021

Communicated by L. Perlovsky

The concept of biological adaptation is closely connected to phenotypic plasticity, which epitomizes the process of phenotypic transition across different states [1] (attractors, in the lexicon of self-organized criticality) [2]. According to Waddington's approach [3] – recently deepened by several theoretical studies [4] – living organisms undergoes numerous symmetry breakings across the differentiating path in response to environmental cues and constraints [5]. Those factors act in a non-instructive manner by selecting (and eventually shaping) the more “appropriate” phenotype (sustained by a corresponding, specific gene-regulatory network) that will maximize system's fitness, thereby adopting the best “adaptive” architecture without requiring genetic (mutation) changes [6]. This perspective help explaining a number of controversial results and – what is more – allows to identify those “tipping point” from which living systems can access into disease states or eventually be committed toward phenotyping “reprogramming” [7]. Yet, attractive as this theoretical framework may be, several hurdles still await compelling resolution. The extensive and remarkable survey by A.N. Gorban et al. [8], provides some fruitful insights by addressing the most critical issues related to this topic.

First, the theoretical framework adopted by Gorban and coworkers requires to rely on a very different concept of causality. It is a shared convention (albeit rarely made explicit in scientific studies) that the kinetics of biochemical reactions and gene transduction (from mRNA production up to protein synthesis) can be described according to the mass-action law. Accordingly, “effects” are viewed as linearly and deterministically determined by the agency of a selected group of genes, whereby complex biological processes are “reduced” to gene activity. However, this paradigm cannot accommodate with an increasing body of observations [9]: to understand biological complexity a shift of attention to patterns and dynamics of causal connections between components is required [10]. To begin with, it is definitely ascertained that genome activity is all but a deterministic process being the gene expression a stochastic process displaying wide changes in space and times [11,12]. Moreover, ordinary differential equations (ODEs) may capture a mass-action relationship only when the reaction kinetics is measured in a homogenous field, where a concentration of reactants and enzymes involves an averaged number of $\sim 6 \times 10^8$. In these terms, the central limit

DOI of original article: <https://doi.org/10.1016/j.plrev.2021.03.001>.

^{*} Corresponding author.

E-mail address: mariano.bizzarri@uniroma1.it (M. Bizzarri).

theorem [13] would almost hold, as stochastic effects and subtle differences in the initial conditions can be “safely” neglected. Unfortunately, within the cell - concentrated in a very restricted (and structured) space of 10–50 μm - there are thousands of different chemical species and (depending on the metazoan specie) ~ 25000 genes, and as many as transcription factors. It is very unlikely that in such conditions each molecular factor could reach a sufficient average value to override environmental noise and fluctuations, thus achieving a deterministic effect [14]. Stochastic effects cannot therefore be ignored as chemical reactions are subject to a non-linear regimen and, consequently, classical ODEs cannot be “linearized” anymore, despite some attempts have been carried out to accommodate randomness to the network framework [15]. Conclusively, models based on continuous modifications of (alleged) key molecular parameters can no longer be advocated to explain processes occurring within complex systems characterized by incessant (chemical and biophysical) fluctuations involving extremely small integer numbers of key and lock molecules. In these systems, the central limit theorem is unsuccessful in describing a number of unfathomable phenomena like bistability in cell fate commitment, paradoxical biological effectiveness of genes, reversibility of cell differentiation, phenotypic plasticity, just to mention a few [16]. The key issue is that randomness is a central feature of the microscopic level, and it is therefore unlikely that microscopic details could help in understanding the logic behind living organisms. Given that causality can be “captured” essentially as a probabilistic concept [17], we should focus on those levels - intermediates between the microscopic and the macroscopic ones - where correlations are maximized. Precisely, this is the mesoscopic level where non-trivial determinism can be grasped by means of aptly application of the Pearson correlation coefficient [18]. Even intuitively, a deterministic relationship is in principle satisfied when the prediction error $r^2 = 1 - E^2/S^2$ approaches 1. This is tantamount to say that the Person's coefficient among two variable X and Y allows to forecast in a quasi-deterministic way the behavior over time of one observable as a function of the other. Gorban et al. furnishes a relevant number of examples in whom the application of correlation graphs has been demonstrated to effectively capture the hidden complexity behind a wide array of different systems, from biology to economy. Noticeably, this approach describes the evolution over time of a living system as a “journey” across a landscape in which the access into new attractors (“valleys”) is allowed after a bifurcation point. It is worth noting that in correspondence of such tipping points, a group of few parameters show a bewildering increase in both variance and reciprocal correlation [19]. As a consequence, the overall connectivity increases, while, in a second step, previous observables that were unaffected at the very beginning, are also “collectively” dragged into the transition to finally enter into a new stable attractor [20]. This remarkable behavior has been vindicated in numerous, different pathophysiological conditions [21], emphasizing that transition toward a new attractor - normal or altered - is not linearly caused by changes in the “functioning” of individual molecules (according to the reductionist paradigm “a gene, a protein, a disease”) but instead by the interaction of a group of highly correlated and intertwined factors. This selected group of molecular factors constitutes a dynamical network biomarker (DNB) [22], which significantly fluctuates immediately before the tipping point has been reached to slow-down during the relaxation phase, when the system accesses a new stable state. Notice that this approach allows to distinguish pre-transition states [23] - i.e., pre-disease from overt disease states - and hence favor recognizing those elements acting as true “causative” factors during the pathophysiological process [24]. This is an outstanding achievement, given that, once the steady state has been reached, the genetic/molecular architecture displayed by the system is usually very different from that of the pre-disease state. However, the system manifests a striking resilience and robustness when challenged by environmental or pharmacological stresses. Destabilization of the previous attractor is then mandatory before a transition could be triggered [25]. Hence, it is noticeably that, by targeting the specific DNB group, one can modify the biological outcome, at least in principle. This assumption has been specifically addressed by studies related to reprogramming of somatic cells [26], or to phenotypic reversion obtained in cancerous cells, in whom the system's stability is challenged by biochemical, biophysical or microenvironmental cues [27–29]. It is worth of note that in cancer cells, the phenotypic transition can be enacted by manipulating the physical forces acting upon the cell, or by adding molecules able in promoting chromatin/cytoskeleton remodeling, hence favoring the rewiring of several networks [30,31]. Thereby, a further conceptual consequence of the model discussed by Gorban et al. is that even stable states (normal as well as pathological) are not irreversibly committed - as usually thought - but can be “reversed” in some specific circumstances. This perspective became attractive in the last decade and has disclosed new opportunities in cancer management [32]. Difficulties arise when we have to identify the DNBs that are associated with the pre-disease state and can forestall the emergence of the tipping point, nonetheless a wealth of sophisticated approaches - high-throughput gene expressions data using multidimensional analytics and deterministic/stochastic middle-out models [33,34] - has been deployed to cope with. Indeed, the network architecture that the system acquires not only depends on the internal agency of genes and proteins as it

is strongly shaped by the environmental conditions in which the experimental model is framed [35]. Gene regulatory networks are differently organized when cell samples are let to growth in 2D or in 3D [36], when the surrounding microenvironment is designed to include stroma parameters (including activated fibroblasts) or engineered into specific topographical profiles [37]. It is therefore obligatory to perform investigations by using proper three-dimensional models able to capture the wealth of the internal/external influences acting upon it. Moreover, phenotypic plasticity and bistability enacted by changes in the biophysical milieu – as those observed in cells cultured in microgravity – can happen without involving appreciable modifications of the overall genome activity [38], and can be promoted by cytoskeleton or molecular pathways remodeling [39].

This is not an easy task, and unfortunately the vast majority of studies still trusts on conventional 2D experimental model. Sad to say, the scientific community involved in biology – largely unaware of the conceptual revolution occurred in physics and thermodynamics [40] – still relies on models and epistemological metaphors that are based on outdated categories, thus showing strong “resilience” in shifting toward a new paradigm. Undoubtedly, “correlation” does not mean “causation”, and correlation studies should be “complemented” by a deeper mechanistic understanding. However, by identifying the relevant hubs and the evolving correlations among them would allow to provide new testable mechanistic hypotheses.

The second half of the Gorban’s review deals with adaption thermodynamics. Starting from Cannon’s and Seyle’s priceless contributions, Gorban and colleagues have strived to identify if some hidden – thermodynamics-like – principles could be advocated in providing a general theoretical framework for the observed transitions across different attractors. It is widely recognized that state transitions in non-living systems are exhaustively described by means of changes in the Gibb’s energy upon the constraint of control parameters. If we adopt the attractor concept to describe the evolution of a living system over space and time, we have to assume that the living entity can evolve from a high energy state toward states characterized by reduced (minimal) energy levels. In principle, entropy should behave inversely: from high entropy values (mirrored by totipotency of stem cells) toward reduced levels, in which specialization strikingly reduces genome totipotence. This hypothetical “entropic landscape” would favor the emergence of convergent solution, i.e., by “attracting” the overall system’s dynamics toward a specific attractor (phenotypic state) with minimal entropy value. The native energy reserve (that can eventually be replenished, as Gorban highlights in the study) would allow the system to cope with stresses and perturbations, finally overcoming stochastic fluctuations and achieving a quasi-deterministic output. Identification of this ‘energy field’ is however problematic, as well as the efforts to reliably measure it. Indeed, this theoretical framework is still in its infancy, despite some promising attempts have been recently proposed [41]. Difficulties lie specifically in identifying a suitable substitute for the Gibb’s potential, notwithstanding some alternatives have been suggested and thoroughly revised by Gorban et al. Furthermore, a landscape in which gradients of energy and entropy (behaving in opposite ways) would be expected to explain the overall transitions – starting from the embryo up to the adult, well-differentiated organism – is a too simplistic hypothesis for dissipative structures for whom classical thermodynamics does not apply. Classical thermodynamics being a theory of states, is not adequate in describing (true) dynamic processes, and different new theoretical perspectives have been proposed according to non-equilibrium thermodynamics [42]. Embracing the theoretical framework for which organisms as dissipative structures governed by nonequilibrium process, will expand our explanatory power beyond a machine-metaphor. However, this is an intricate matter that deserves a proper treatise in its own. Yet, a major drawback upon which I wish to draw the attention is the need for a measurable parameter that could be considered a reliable substitute for “cell entropy”. Promising results have been recently proposed by considering signaling entropy, which quantifies the degree of uncertainty (promiscuity) of a cell’s gene expression levels in the context of a cellular interaction network [43]. This approach provides a more accurate and robust potency estimate than other entropy-based measures and allows recognizing different phenotypes within an apparently homogenous population, and how this heterogeneity within single-cell populations is regulated [44]. Therefore, this measure can furnish a valuable proxy of the differentiation potency and plasticity of cells and tissues [45].

The accurate assay from Gorban et al. has the merit to having raised these issues by clearly establishing where we are and where we should go to assemble the very basic bricks of a new theoretical framework for understanding biological processes.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] Goldberg A, Allis CD, Bernstein E. Epigenetics: a landscape takes shape. *Cell* 2007;128:635–8. <https://doi.org/10.1016/j.cell.2007.02.006>.
- [2] Bak P. *How nature works: the science of self-organized criticality*. Springer Science & Business Media; 2013.
- [3] Waddington CH. *The strategy of the genes: A discussion of some aspects of theoretical biology*. London: Allen & Unwin; 1957.
- [4] Longo G, Montévil M. From physics to biology by extending criticality and symmetry breakings. *Prog Biophys Mol Biol* 2011;106(2):340–7. <https://doi.org/10.1016/j.pbiomolbio.2011.03.005>.
- [5] Bizzarri M, Giuliani A, Minini M, Monti N, Cucina A. Constraints shape cell function and morphology by canalizing the developmental path along the Waddington's landscape. *BioEssays* 2020;42(4):e1900108. <https://doi.org/10.1002/bies.201900108>.
- [6] West-Eberhard MJ. Phenotypic accommodation: adaptive innovation due to developmental plasticity. *J Exp Zool, Part B Mol Dev Evol* 2005;304(6):610–8. <https://doi.org/10.1002/jez.b.21071>.
- [7] Bizzarri M, Palombo A, Cucina A. Theoretical aspects of systems biology. *Prog Biophys Mol Biol* 2013;112(1–2):33–43. <https://doi.org/10.1016/j.pbiomolbio.2013.03.019>.
- [8] Gorban AN, Tyukina TA, Pokidysheva LI, et al. Dynamic and thermodynamic models of adaptation. *Phys Life Rev* 2021;37:17–64.
- [9] Paldi A. Stochastic or deterministic? That is the question. *Org: J Biol Sci* 2020;4(1):77–9. <https://doi.org/10.13133/2532-5876/16963>.
- [10] Bizzarri M, Brash DE, Briscoe J, Grieneisen VA, Stern CD, Levin M. Publisher correction: a call for a better understanding of causation in cell biology. *Nat Rev Mol Cell Biol* 2019;20(6):384. <https://doi.org/10.1038/s41580-019-0130-6>. Erratum for: *Nat Rev Mol Cell Biol* 2019;20(5):261–2.
- [11] Kupiec JJ. A probabilistic theory for cell differentiation, embryonic mortality and DNA C-value paradox. *Org: J Biol Sci* 2020;4(1):80–5. <https://doi.org/10.13133/2532-5876/16964>.
- [12] Eldar A, Elowitz M. Functional roles for noise in genetic circuits. *Nature* 2010;467:167–73. <https://doi.org/10.1038/nature09326>.
- [13] Montgomery DC, Runger GC. *Applied statistics and probability for engineers*. 6th ed. Wiley. ISBN 9781118539712, 2014. p. 241.
- [14] Yoshikawa K. Field hypothesis on the self-regulation of gene expression. *J Biol Phys* 2002;28:701–12. <https://doi.org/10.1023/A:1021251125101>.
- [15] McAdams HH, Arkin A. Stochastic mechanisms in gene expression. *Proc Natl Acad Sci USA* 1997;94(3):814–9. <https://doi.org/10.1073/pnas.94.3.814>.
- [16] Bizzarri M, Naimark O, Nieto-Villar J, Fedeli V, Giuliani A. Complexity in biological organization: deconstruction (and subsequent restating) of key concepts. *Entropy* 2020;22(8):885. <https://doi.org/10.3390/e22080885>.
- [17] Albantakis L, Marshall W, Hoel E, Tononi G. What caused what? A quantitative account of actual causation using dynamical causal networks. *Entropy* 2019;21(5):459. <https://doi.org/10.3390/e21050459>.
- [18] Laughlin RB, Pines D, Schmalian J, Stojković BP, Wolynes P. The middle way. *Proc Natl Acad Sci USA* 2000;97(1):32–7. <https://doi.org/10.1073/pnas.97.1.32>.
- [19] Tsuchiya M, Hashimoto M, Takenaka Y, Motoike IN, Yoshikawa K. Global genetic response in a cancer cell: self-organized coherent expression dynamics. *PLoS ONE* 2014;9(5):e97411. <https://doi.org/10.1371/journal.pone.0097411>. Erratum in: *PLoS ONE* 2014;9(8):e105491.
- [20] Mojtahedi M, Skupin A, Zhou J, Castaño IG, Leong-Quong RY, Chang H, et al. Fate decision as high-dimensional critical state transition. *PLoS Biol* 2016;14(12):e2000640. <https://doi.org/10.1371/journal.pbio.2000640>.
- [21] Bose I, Pal M. Criticality in cell differentiation. *J Biosci* 2017;42(4):683–93. <https://doi.org/10.1007/s12038-017-9721-6>.
- [22] Chen L, Liu R, Liu ZP, Li M, Aihara K. Detecting early-warning signals for sudden deterioration of complex diseases by dynamical network biomarkers. *Sci Rep* 2012;2:1–8.
- [23] Koizumi K, Oku M, Hayashi S, Inujima A, Shibahara N, Chen L, et al. Identifying pre-disease signals before metabolic syndrome in mice by dynamical network biomarkers. *Sci Rep* 2019;9:1–11. <https://doi.org/10.1038/s41598-019-45119-w>.
- [24] Liu X, Chang X, Leng S, Tang H, Aihara K, Chen L. Detection for disease tipping points by landscape dynamic network biomarkers. *Nat Sci Rev* 2019;6:775–85. <https://doi.org/10.1093/nsr/nwy162>.
- [25] Paldi A. Random walk across the epigenetic landscape. In: *Phenotypic switching implications in biology and medicine*. Academic Press; 2020. p. 53–76.
- [26] Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 2006;126:663–76. <https://doi.org/10.1016/j.cell.2006.07.024>.
- [27] Kenny PA, Bissell MJ. Tumor reversion: correction of malignant behavior by microenvironmental cues. *Int J Cancer* 2003;107:688–95. <https://doi.org/10.1002/ijc.11491>.
- [28] Downing TL, Soto J, Morez C, Houssin T, Fritz A, Yuan F, et al. Biophysical regulation of epigenetic state and cell reprogramming. *Nat Mater* 2013;12(12):1154–62. <https://doi.org/10.1038/nmat3777>.
- [29] Proietti S, Cucina A, Pensotti A, Fuso A, Marchese C, Nicolini A, et al. Tumor reversion and embryo morphogenetic factors. *Semin Cancer Biol* 2020. S1044-579X(20)30194-2. <https://doi.org/10.1016/j.semcancer.2020.09.005> [in press].
- [30] Lavrovsky VA, Guvakova MA, Lavrovsky YV. High frequency of tumour cell reversion to non-tumorigenic phenotype. *Eur J Cancer* 1992;28(1):17–21. [https://doi.org/10.1016/0959-8049\(92\)90375-c](https://doi.org/10.1016/0959-8049(92)90375-c).

- [31] Gaspar-Maia A, Alajem A, Meshorer E, Ramalho-Santos M. Open chromatin in pluripotency and reprogramming. *Nat Rev Mol Cell Biol* 2011;12:36–47. <https://doi.org/10.1038/nrm3036>.
- [32] Powers S, Pollack RE. Inducing stable reversion to achieve cancer control. *Nat Rev Cancer* 2016;16(4):266–70. <https://doi.org/10.1038/nrc.2016.12>.
- [33] Simeoni O, Piras V, Tomita M, Selvarajoo K. Tracking global gene expression responses in T cell differentiation. *Gene* 2015;569(2):259–66. <https://doi.org/10.1016/j.gene.2015.05.061>.
- [34] Deveaux W, Hayashi K, Selvarajoo K. Defining rules for cancer cell proliferation in TRAIL stimulation. *npj Syst Biol Appl* 2019;5:5. <https://doi.org/10.1038/s41540-019-0084-5>.
- [35] Bloom AB, Zaman MH. Influence of the microenvironment on cell fate determination and migration. *Physiol Genomics* 2014;46(9):309–14. <https://doi.org/10.1152/physiolgenomics.00170.2013>.
- [36] Fontoura JC, Viezzer C, Dos Santos FG, Ligabue RA, Weinlich R, Puga RD, et al. Comparison of 2D and 3D cell culture models for cell growth, gene expression and drug resistance. *Mater Sci Eng C, Mater Biol Appl* 2020;107:110264. <https://doi.org/10.1016/j.msec.2019.110264>.
- [37] Yoo J, Kim J, Baek S, Park Y, Im H, Kim J. Cell reprogramming into the pluripotent state using graphene based substrates. *Biomaterials* 2014;35(29):8321–9. <https://doi.org/10.1016/j.biomaterials.2014.05.096>.
- [38] Po A, Giuliani A, Masiello MG, Cucina A, Catizone A, Ricci G, et al. Phenotypic transitions enacted by simulated microgravity do not alter coherence in gene transcription profile. *npj Microgravity* 2019;5:27. <https://doi.org/10.1038/s41526-019-0088-x>.
- [39] Dinicola S, Fabrizi G, Masiello MG, Proietti S, Palombo A, Minini M, et al. Inositol induces mesenchymal-epithelial reversion in breast cancer cells through cytoskeleton rearrangement. *Exp Cell Res* 2016;345(1):37–50. <https://doi.org/10.1016/j.yexcr.2016.05.007>.
- [40] Prigogine I, Stengers I. *Order out of chaos: Man's new dialogue with nature*. Flamingo. ISBN 0-00-654115-1, 1984.
- [41] Longo G, Montévil M. *Perspectives on organisms*. Berlin Heidelberg: Springer-Verlag; 2014.
- [42] Kondepudi DK, De Bari B, Dixon JA. Dissipative structures, organisms and evolution. *Entropy* 2020;22(11):1305. <https://doi.org/10.3390/e22111305>.
- [43] Teschendorff AE, Sollich P, Kuehn R. Signalling entropy: a novel network-theoretical framework for systems analysis and interpretation of functional omic data. *Methods* 2014;67(3):282–93. <https://doi.org/10.1016/j.ymeth.2014.03.013>.
- [44] Teschendorff AE, Enver T. Single-cell entropy for accurate estimation of differentiation potency from a cell's transcriptome. *Nat Commun* 2017;8:15599. <https://doi.org/10.1038/ncomms15599>.
- [45] Nijman SMB. Perturbation-driven entropy as a source of cancer cell heterogeneity. *Trends Cancer* 2020;6(6):454–61. <https://doi.org/10.1016/j.trecan.2020.02.016>.

Comment

Ensemble of correlation, parenclitic and synolitic graphs as a tool to detect universal changes in complex biological systems

Comment on “Dynamic and thermodynamic models of adaptation” by A.N. Gorban et al.

Tatiana Nazarenko^a, Oleg Blyuss^{a,b,c,g}, Harry Whitwell^{c,d,e,f}, Alexey Zaikin^{a,e,f,*}^a Department of Mathematics and Institute for Women's Health, University College London, London, UK^b School of Physics, Astronomy and Mathematics, University of Hertfordshire, Harfield, UK^c National Phenome Centre and Imperial Clinical Phenotyping Centre, Department of Metabolism, Digestion and Reproduction, IRDB Building, Imperial College London, Hammersmith Campus, London, W12 0NN, UK^d Section of Bioanalytical Chemistry, Division of Systems Medicine, Department of Metabolism, Digestion, Imperial College London, South Kensington Campus, London, SW7 2AZ, UK^e Lobachevsky State University of Nizhny Novgorod, Nizhny Novgorod, Russia^f Centre for Analysis of Complex Systems, Sechenov First Moscow State Medical University (Sechenov University), Moscow, Russia^g Department of Pediatrics and Pediatric Infectious Diseases, Institute of Child's Health, Sechenov First Moscow State Medical University (Sechenov University), Moscow, Russia

Received 24 May 2021; accepted 24 May 2021

Available online 28 May 2021

Communicated by L. Perlovsky

Keywords: Adaptation; Universality; Thermodynamics; Complexity; Correlation graphs; Parenclitic graphs

Complexity is a natural feature of biological systems. For example, a human can be characterised by a vast amount of information, packed in a super-network that is, in principle, a network of networks [1]. For each human cell, the total amount of DNA, if measured in bases, has an informational length of the 3.2 Gb. However, in total, we have approximately 10^{12} cells, and in each of them, even neglecting mosaic mutations, the genotype information is multiplied by the epigenome, transcriptome, proteome and metabolome profiles, which are organized in a huge and very dynamic connectome. Each tissue can be represented by a graph or network and connected to other networks, resulting in a super-network containing a huge number of features describing a complex biological system. When the biological system ages, or approaches a diseased state, the bodies adaptation to maintaining homeostasis is reflected in the changes of this super-network.

Due to the fascinating progress in multi-omic experimental techniques, we are getting access to increasingly huge quantities of biological data. But it is clear, that the amount of information and, more importantly, the complexity of

DOI of original article: <https://doi.org/10.1016/j.plrev.2021.03.001>.

* Corresponding author.

E-mail address: alexey.zaikin@ucl.ac.uk (A. Zaikin).

its organisation, exceeds our analytical capability. The situation is even more complicated, because, for example, as noted by Gorban et al. in the analysis of cancer omics data, underlying changes may occur in different places of the network. A possible solution to the challenging problem of identifying biomarkers and disease-precursors is in graph theory, using well-described topological descriptors to identify changes in network structure that represent disease. In this sense, the “shape” of the network is used, rather than specific molecular biomarkers. However, a real breakthrough can only occur if we are able to define the behaviour of the networks to a relevant stimulus (e.g. disease), in a generic way, for example, based on the fundamental laws of thermodynamics.

Representing high-dimensional biological data in the form of a graph and linking features by biological and thermodynamic laws seems to be a very promising approach to deal with overwhelming complexity of biological systems. However, one can utilise this approach only if we have information about how features and attributes are connected biologically. If we do not have such information, we must use more sophisticated methods to represent a complex system in the form of the graph in high dimensional space. Such methodology is provided by correlation graphs which are carefully reviewed in fascinating and inciteful detail in the review by Gorban and co-authors [2]. This paper considers two types of correlation graph. Firstly, where edges correspond to the correlations between attributes of different people, with populations represented as a network and secondly, where edges correspond to different time intervals between data belonging to the same person. Most importantly, the authors reviewed and discussed a link between changes in the topology and laws of a general theory of adaptation and thermodynamics. In particular, it is discussed that preceding a network crash or critical state, there is a detectable redistribution of resources within the system resulting in the strengthening of correlations, breaking of topological symmetry and increase of fluctuation variances.

In this comment we would like to draw attention to alternative methods to represent high dimensional data in the form of the graph if *a-priori* we do not have established connections. There are several other methods that we will briefly mention here.

First of all, correlation-prediction graphs can be used as a marker of survival and have been constructed to represent the gene methylation profiles of individuals [5,6]. Secondly, there is an algorithm, first described by Zanin and Bocaletti, able to establish links between parameters/nodes without any *a-priori* knowledge of their interactions [3] using residual distances from linear regression models constructed between every pair of analytes to construct a graph. They termed this approach a “parenclitic” network representation, from the Greek term for “deviation”. Parenclitic networks have been successfully applied to problems of the detection of key genes and metabolites in different diseases [3,7–9] see [10] for a review and [11] for a discussion of applications for brain research. In [12] we have applied this methodology to implement machine learning classification to identify signatures of cancer development from human DNA methylation data. Thirdly, based on the understanding that the interactions of two features (at least in biological systems) often cannot be described by a linear model, it was proposed to use 2-dimensional kernel density estimation (2DKDE) to model the control distribution [4]. Finally, in [13] we have introduced a variation of parenclitic networks, that can be called synolytic from the Greek word for “ensemble”. In principle, these networks can be considered an ensemble of classifiers in a graph form and thus are a kind of correlation network where the correlation is in the changes between two classes (e.g. disease and non disease). The difference to original parenclitic networks is visualized in the Fig. 1. These networks have been successfully used to detect age related trajectories in Down’s syndrome [13] and for prediction of survival for severely ill Covid-19 patients [14].

All these network approaches could be described as variations of correlation networks with different levels of dimensionality. In other words, an ensemble of different correlation graphs can be used to characterise the same complex biological system. This ensemble of graphs as a classifier can provide us with more efficient diagnostic methods than single classifiers. The motivation behind combining the classifiers as an ensemble is based on the classical Condorcet’s Jury theorem stating that the probability of a given group of individuals (classifiers) arriving at a correct decision (correct class label) based on a majority vote, increases with the total number of voters. The theorem requires that one of the two outcomes should be correct, and each voter has an independent probability $p > 1/2$ of voting for the correct decision. The assumption that the votes are independent is not always easy to satisfy but a common approach is to train the classifiers on different subsets of the data. A further advantage of combining classifiers into an ensemble is a gain in robustness to noise that may accompany the data.

Designing an optimal aggregation scheme for combining multiple classifiers remains challenging [15] despite a number of methodologies suggested. It is very promising to exploit these different methodologies to construct correlation graphs and use them as an ensemble to develop new diagnostic tools. It would be especially valuable to

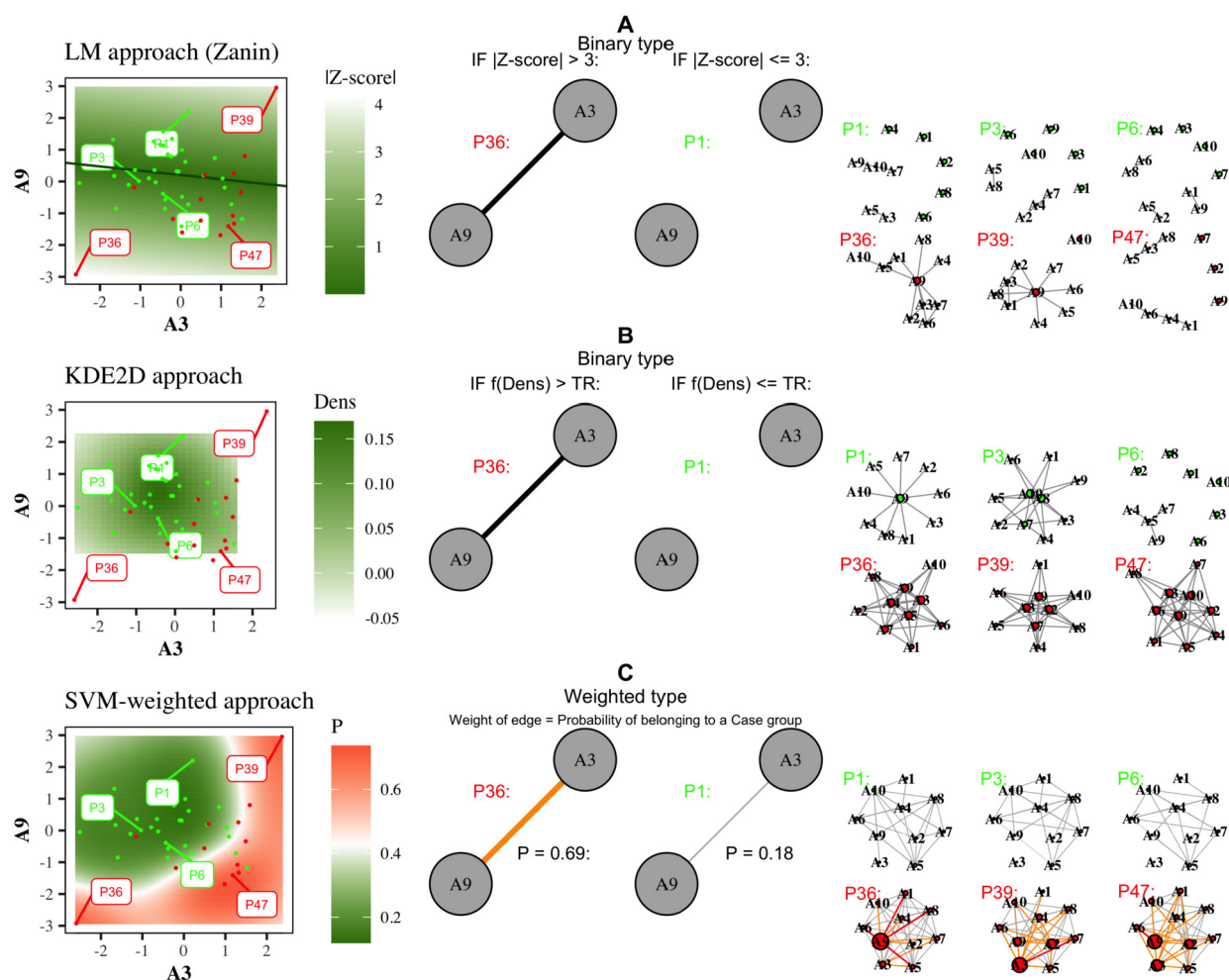


Fig. 1. Parenclitic and synolytic networks as particular types of a correlation graph: (A) (*Original parenclitic network* [3]). This is a network where only the control group was considered as the basis for determining the normal state on the plane of two signs (based on control group the linear regression was built, the deviation of the control points from it was calculated and the distribution of such deviations was constructed). For each new sample, the edge weight was first determined as the absolute value of z-score, and then binarized (if z-score is greater than 3, then the edge is present in the sample network, otherwise there is no edge). (B) (*KDE-parenclitic network* [4]). This is a network in which again only the control group was considered as the basis for determining the normal state on the plane of two signs (for the control group, the KDE2D density was built, then a function was calculated that converts the density values into an analogue distance (so that the points located in the area of the highest density have the minimum weight); the distance outside the grid was continued (for more details, see [4]). For each new sample, the edge weight was first determined as the value of the entered distance, and then converted to binary form (if the distances are greater than a threshold, then the edge is present in the sample network, otherwise there is no edge). (C) (*SVM-weighted synolytic network*). This is a network in which both groups participate in defining normal and abnormal states. On the plane of two features, with the help of the radial SVM, the best boundary separating the classes is drawn. Automatically, each point in such a model gets a value for the probability of belonging to each class. For each new sample, the edge weight is determined as the probability of belonging to a group of cases.

link topological features of these graphs with universal phenomena, explainable with thermodynamic laws as it was discussed in [2], thus providing a generalisation that can greatly benefit their ability to predict disease.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

We thank support from MRC MR/R02524X/1 and the Ministry of Science and Higher Education of the Russian Federation (project no. 075-15- 2020-808).

References

- [1] Whitwell HJ, Bacalini MG, Blyuss O, Chen S, Garagnani P, Gordleeva SY, et al. The human body as a super network: digital methods to analyze the propagation of aging. *Front Aging Neurosci* 2020;12:136.
- [2] Gorban A, Tyukina T, Pokidysheva L, Smirnova E. Dynamic and thermodynamic models of adaptation. *Phys Life Rev* 2021;37:17–64. <https://doi.org/10.1016/j.plrev.2021.03.001>.
- [3] Zanin M, Boccaletti S. Complex networks analysis of obstructive nephropathy data. *Chaos, Interdiscip. J. Nonlinear Sci.* 2011;21:033103.
- [4] Whitwell HJ, Blyuss O, Menon U, Timms JF, Zaikin A. Parenclitic networks for predicting ovarian cancer. *Oncotarget* 2018;9:22717–26. <https://doi.org/10.18632/oncotarget.25216>.
- [5] Bartlett TE, Zaikin A. Detection of epigenomic network community oncomarkers. *Ann Appl Stat* 2016;10:1373–96. <https://doi.org/10.1214/16-AOAS939>.
- [6] Bartlett TE, Olhede SC, Zaikin A. A dna methylation network interaction measure, and detection of network oncomarkers. *PLoS ONE* 2014;9:e84573.
- [7] Zanin M, Menasalvas E, Sousa PA, Boccaletti S. Preprocessing and analyzing genetic data with complex networks: an application to obstructive nephropathy. *Netw Heterog Media* 2012;7:473.
- [8] Zanin M, Menasalvas E, Boccaletti S, Sousa P. Feature selection in the reconstruction of complex network representations of spectral data. *PLoS ONE* 2013;8:e72045.
- [9] Zanin M, Papo D, Solís JLG, Espinosa JCM, Frausto-Reyes C, Anda PP, et al. Knowledge discovery in spectral data by means of complex networks. *Metabolites* 2013;3:155–67.
- [10] Zanin M, Papo D, Sousa PA, Menasalvas E, Nicchi A, Kubik E, et al. Combining complex networks and data mining: why and how. *Phys Rep* 2016;635:1–44.
- [11] Papo D, Buldú JM, Boccaletti S, Bullmore ET. Introduction: complex network theory and the brain, *philosophical transactions. Biol Sci* 2014;369:1–7.
- [12] Karsakov A, Bartlett T, Ryblov A, Meyerov I, Ivanchenko M, Zaikin A. Parenclitic network analysis of methylation data for cancer identification. *PLoS ONE* 2017;12:e0169661.
- [13] Krivonosov M, Nazarenko T, Bacalini M, Franceschi C, Zaikin A, Ivanchenko M. DNA methylation changes with age as a complex system: a parenclitic network approach to a family-based cohort of patients with down syndrome. *bioRxiv* <https://doi.org/10.1101/2020.03.10.986505>.
- [14] Demichev V, Tober-Lau P, Nazarenko T, Thibeault C, Whitwell H, Lemke O, et al. A time-resolved proteomic and diagnostic map characterizes Covid-19 disease progression and predicts outcome. *medRxiv* <https://doi.org/10.1101/2020.11.09.20228015>.
- [15] Kuncheva LI. Combining pattern classifiers: methods and algorithms. John Wiley & Sons; 2014.

Comment

Network models of biological adaptation at the molecular scale

Comment on “Dynamic and thermodynamic models of adaptation”

by A.N. Gorban et al.

Luisa Di Paola^{a,*}, David M. Leitner^b^a *Unit of Chemical-physics Fundamentals in Chemical Engineering, Department of Engineering, Università Campus Bio-Medico di Roma, via Álvaro del Portillo 21, Rome, 00128, Italy*^b *Department of Chemistry and Chemical Physics Program, University of Nevada, Reno, 89557, NV, USA*

Received 23 May 2021; accepted 24 May 2021

Available online 28 May 2021

Communicated by L. Perlovsky

The inspiring work of Gorban and coworkers [1] highlights how correlation graphs explain connections among a variety of processes from scales of individual organisms to societies. Networks, and correlation graphs in particular, capture relationships among the myriad elements of complex systems [2] (e.g., the structure of social networks - whether digital or not - frames the scale of relevant events in society). The intriguing aspect of networks is their ability to describe very different systems, from the life sciences to finance to technology, with a general language.

Being a matter of definition, the elements in a life network (nodes) span scales from regions of molecules to whole societies. We focus here on a smaller limit of observation with respect to the kinds of systems that were discussed in Ref. [1]: the molecular scale of life, embodied by proteins, a basic element of living systems, albeit one with properties of the non-biological world. Protein science lies at the interface of physics, chemistry and biology [3]; the functional feature addressing their role as biological elements is allostery, the molecular machinery behind homeostatic processes.

The success of structural networks relies on the right choice of scale of interaction they account for. In the case of molecular graphs, representing the structural formula of organic molecules, network links represent π orbitals, responsible for many molecular properties [4]. As such, Protein Contact Networks (PCNs) represent the intramolecular interactions (noncovalent, with a prevalence of hydrophobic interactions) responsible for protein dynamic adaptation to its environment. In a sense, PCNs represent a class of causal correlation graphs, since the intramolecular interactions are a cause of correlations between elements (residues) in the protein dynamics and function.

Allosteric regulation, the control of protein activity by processes, such as ligand binding, that occur at sites often far removed, plays a critical role in cell function. Network analysis has identified regions of proteins and systems of biomolecules involved in signaling contributing to allosteric regulation, including “hot spot” residues that mediate the signal. How protein function, including allostery, adapts to stress such as mutation is currently a major area of study. Recent work [5] suggests that proteins adapt to mutations by developing new allosteric networks satisfying thermo-

DOI of original article: <https://doi.org/10.1016/j.plrev.2021.03.001>.

* Corresponding author.

E-mail address: l.dipaola@unicampus.it (L. Di Paola).

dynamic conditions for allostery. Residues important for structural integrity are highly conserved, while residues that appear as allosteric hot spots are not. There thus appears to be evolutionary pressure for structural stability, whereas allosteric networks exhibit plasticity; signaling does not require a unique pathway and is adaptable to mutation. Upon mutation a “rewiring” of allosteric networks occurs, analogous to adaptation of genetic networks to some diseases [1].

Networks facilitate identifying allosteric regions in protein molecules and complexes through structural and dynamic descriptions. They help to pinpoint residues, or hot spots, that contribute to signal transmission and mechanisms that underlie allostery. Networks have been defined and analyzed both in terms of structural as well as dynamic properties.

Dynamic networks indicate signaling pathways in proteins and protein complexes that facilitate energy transport between binding sites [6–11]. For example, energy transport networks of the homodimeric hemoglobin from *Scapharca inaequivalvis* (HbI), identified by molecular dynamics simulations [12], reveal regions of the protein including water molecules that mediate allosteric transitions. Those networks are altered by mutation, reconfigured to facilitate cooperative ligand binding, albeit by a different mechanism. The dynamic networks obtained in the computational studies of HbI and mutants [12,13] indicate regions that mediate cooperativity of ligand binding previously identified experimentally [14,15].

Those results suggest that allostery adapts to mutation by at least some reorganization of networks, a general picture put forth and supported by a recent study involving mutational scanning of a bacterial tetracycline repressor [5], indicating rearrangement of networks for allostery in response to a variety of mutations. The networks that emerge from adaptation to mutation have to satisfy the thermodynamic constraints for allostery. Small structural adjustments can lead to significant changes in dynamics throughout the protein [16], and corresponding entropy, which can facilitate ligand binding and allostery [17].

Protein contact networks (PCNs) are structural networks whose elements are the protein residues and links between nodes represent the relevant intramolecular interactions (noncovalent) in protein structures [18]. They allow to identify residues active in allosteric regulation through partition of the protein molecules in functional domains [19,20]. For instance, in enzyme systems, PCNs let emerge functional domains and the border regions as responsible for transmission of signals between domains [21].

Recently, we have applied the PCN methodology to verify the presence of allosteric regions in the complex of the spike protein of SARS-CoV2 with the human receptor ACE2. We have followed an integrated dynamic/topological approach comparing results with those obtained through a method that predicts intersubunit affinity (SEPAS) as well as the Elastic Network Model (ENM). There was general agreement between all the methods applied. PCNs have identified residues of an allosteric modulation region (AMR) in charge of allosteric communication between the binding interface and the rest of the protein molecule [20].

All in all, we can say the PCNs and dynamical networks in protein molecular systems represent the molecular scale of (causal) correlation graphs, falling in the same general framework of the work of Gorban [1].

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Support from NSF CHE-1854271 (to DML) is gratefully acknowledged.

References

- [1] Gorban AN, Tyukina TA, Pokidysheva LI, Smirnova EV. Dynamic and thermodynamic models of adaptation. *Phys Life Rev* 2021;37:17–64. <https://doi.org/10.1016/j.plrev.2021.03.001>.
- [2] Barabási A-L. Taming complexity. *Nat Phys* Jan 2012;8:14–6.
- [3] Shakhnovich E. Protein folding thermodynamics and dynamics: where physics, chemistry, and biology meet. *Chem Rev* 2006;106(5):1559–88. <https://doi.org/10.1021/cr040425u>.
- [4] Krivka P, Mallion R, Trinajstić N. Chemical graph theory. *J Mol Struct, Theochem* 2002;164(3–4):363–77. [https://doi.org/10.1016/0166-1280\(88\)80157-x](https://doi.org/10.1016/0166-1280(88)80157-x).

- [5] Leander M, Yuan Y, Meger A, Cui Q, Raman S. Functional plasticity and evolutionary adaptation of allosteric regulation. *Proc Natl Acad Sci USA* 2020;117(41):25445–54. <https://doi.org/10.1073/pnas.2002613117>.
- [6] Leitner DM. Energy flow in proteins. *Annu Rev Phys Chem* 2008;59:233–59. <https://doi.org/10.1146/annurev.physchem.59.032607.093606>.
- [7] Smock RG, Gierasch LM. Sending signals dynamically. *Science* 2009;324(5924):198–203. <https://doi.org/10.1126/science.1169377>.
- [8] Sethi A, Eargle J, Black AA, Luthy-Schulten Z. Dynamical networks in tRNA: protein complexes. *Proc Natl Acad Sci USA* 2009;106(16):6620–5. <https://doi.org/10.1073/pnas.0810961106>.
- [9] Feher VA, Durrant JD, Wart ATV, Amaro RE. Computational approaches to mapping allosteric pathways. *Curr Opin Struct Biol* 2014;25:98–103. <https://doi.org/10.1016/j.sbi.2014.02.004>.
- [10] Gheeraert A, Pacini L, Batista VS, Vuillon L, Lesieur C, Rivalta I. Exploring allosteric pathways of a v-type enzyme with dynamical perturbation networks. *J Phys Chem B* 2019;123:3452–61. <https://doi.org/10.1021/acs.jpcb.9b01294>.
- [11] Ota K, Yamato T. Energy exchange network model demonstrates protein allosteric transition: an application to an oxygen sensor protein. *J Phys Chem B* 2019;123(4):768–75. <https://doi.org/10.1021/acs.jpcb.8b10489>.
- [12] Leitner DM, Hyeon C, Reid KM. Water-mediated biomolecular dynamics and allostery. *J Chem Phys* 2020;152:240901. <https://doi.org/10.1063/5.0011392>.
- [13] Gnanasekaran R, Agbo JK, Leitner DM. Communication maps computed for homodimeric hemoglobin: computational study of water-mediated energy transport in proteins. *J Chem Phys* 2011;135:065103. <https://doi.org/10.1063/1.3623423>.
- [14] Royer WE, Pardanani A, Gibson QH, Peterson ES, Friedman JM. Ordered water molecules as key allosteric mediators in a cooperative dimeric hemoglobin. *Proc Natl Acad Sci USA* 1996;93:14526–31. <https://doi.org/10.1073/pnas.93.25.14526>.
- [15] Pardanani A, Gambacurta A, Ascoli F, Royer WE. Mutational destabilization of the critical interface water cluster in scapharca dimeric hemoglobin: structural basis for altered allosteric activity. *J Mol Biol* 1998;284:729–39. <https://doi.org/10.1006/jmbi.1998.2195>.
- [16] Reid KM, Yu X, Leitner DM. Change in vibrational entropy with change in protein volume estimated with mode Grüneisen parameters. *J Chem Phys* 2021;154:055102. <https://doi.org/10.1063/5.0039175>.
- [17] Bozovic O, Zanolini C, Gulzar A, Jankovic B, Buhrke D, Post M, et al. Real-time observation of ligand-induced allosteric transitions in a PDZ domain. *Proc Natl Acad Sci USA* 2020;117(42):26031–9. <https://doi.org/10.1073/pnas.2012999117>.
- [18] Di Paola L, De Ruvo M, Paci P, Santoni D, Giuliani A. Protein contact networks: an emerging paradigm in chemistry. *Chem Rev* 2013;113(3):1598–613. <https://doi.org/10.1021/cr3002356>.
- [19] Di Paola L, Giuliani A. Protein contact network topology: a natural language for allostery. *Curr Opin Struct Biol* 2015;31:43–8. <https://doi.org/10.1016/j.sbi.2015.03.001>.
- [20] Di Paola L, Hadi-Alijanvand H, Song X, Hu G, Giuliani A. The discovery of a putative allosteric site in the SARS-CoV-2 spike protein using an integrated structural/dynamic approach. *J Proteome Res* 2020;19(11):4576–86. <https://doi.org/10.1021/acs.jproteome.0c00273>.
- [21] Cimini S, Di Paola L, Giuliani A, Ridolfi A, De Gara L. GH₃₂ family activity: a topological approach through protein contact networks. *Plant Mol Biol* 2016;92(4–5):401–10. <https://doi.org/10.1007/s11103-016-0515-2>.

Comment

Exercise science perspective

Comment on “Dynamic and thermodynamic models of adaptation” by Alexander N. Gorban et al.

Ecaterina Vasenina^a, Ryo Kataoka^a, Jeremy P. Loenneke^b, Samuel L. Buckner^{a,*}^a *USF Muscle Lab, Exercise Science Program, University of South Florida, Tampa, FL, United States of America*^b *Department of Health, Exercise Science, and Recreation Management, Kevser Ermin Applied Physiology Laboratory, The University of Mississippi, University, MS, United States of America*

Received 19 May 2021; accepted 19 May 2021

Available online 23 May 2021

Communicated by L. Perlovsky

Within the field of exercise science, Hans Selye's general adaptation syndrome (GAS) has risen to prominence, as it is believed to help explain the adaptative response to exercise [1]. Interestingly, this has occurred with little mention or acknowledgment of the concept of adaptation energy (AE), which Hans Selye used as a physiological rationale to explain his experimental observations [2]. Within a short communication in the *Journal of Track Technique* one can find only a brief mention of AE. Fred Wilt wrote:

“Selye has indicated that the body has a finite quantity of adaptation energy which it can transfer from one activity to another. Prolonged or chronic exposure of the body to one specific stress will increase resistance to that stress, but at the expense of reducing non-specific resistance to other stressors. Any stress draws from the body's finite quantity of adaptation energy. Those stressors to which the body is less well adapted will make the greatest demand on the body's store of adaptation energy, and be most apt to hasten exhaustion reactions” [3].

Following Wilt's brief commentary [3], the GAS was adopted by exercise science (at least in name); but replaced the phenomena observed by Hans Selye with different phenomena that occurred following exercise [4]. For example, Selye's observation of thymus involution in rats (amongst other structural changes) following exposure to a toxic stressor (alarm phase) [2] was compared to the muscle soreness and decreased performance experienced in the hours and days following an exercise bout [3]. During the second phase of the GAS (resistance), Selye noted a resistance to a given toxic stressor, whereby the thymus (and other organ changes) would return to baseline values/sizes despite continued exposure to the damaging stimulus [2]. Within exercise science this was compared to the period during which individuals experienced favorable adaptations such as increased muscle size and strength from training [4]. Finally, Selye observed that all toxic stressors would eventually lead to death of the organism if the stress persisted (exhaustion) [2]. Within the field of exercise science this has been associated with overtraining (i.e. prolonged performance decrements) [4]. Altogether, these interpretations of the GAS within exercise science are brought together in the su-

DOI of original article: <https://doi.org/10.1016/j.plrev.2021.03.001>.

* Corresponding author at: 4202 E. Fowler Ave. PED 214 Tampa, FL, 33620-8600, United States of America.

E-mail address: slbuckner@usf.edu (S.L. Buckner).

percompensation hypothesis of adaptation. This hypothesis suggests that a decrease in a fitness or performance metric (alarm phase) will stimulate an increase in the robustness of that fitness or performance metric (resistance phase). The stress of training is managed in an attempt to avoid overtraining (exhaustion phase). Within this model, the recovery does not just reach baseline, but overshoots (termed supercompensation) [5]. Although widely accepted, this application of Selye's GAS to voluntary exercise in humans has been criticized, with some authors suggesting that the stress observed in Hans Selye's experiments on the GAS is remarkably different than the stress that can be observed in humans following voluntary exercise [6]. The supercompensation hypothesis is proposed as a mechanistic explanation for adaptation, however, critics note that this idea is more of a time-course than a mechanism [5]. For example, it has been demonstrated that fatigue can be largely explained by both the short-term accumulation of metabolic byproducts and/or the depletion of energy substrates [7]. In addition, muscle growth following training appears to involve the conversion of mechanical signals to molecular cascades which result in the activation or repression of pathways that alter gene expression [8] and protein synthesis [9]. The primary disconnect appears to be exercise science's direct physical interpretation of stages of the GAS. Moreover, when the GAS is applied, it is done without acknowledging AE.

When examining the concept of AE, it is challenging to determine the relevance it may have regarding voluntary exercise in humans [10]. However, the top down approach proposed by Gorban et al. [11,12] may allow for new incorporations and evolutions of the concept of AE within exercise science. For example, the fitness fatigue model of adaption suggests that individuals accumulate both fitness and fatigue over time [13]. Although not completely understood, exercise science provides several mechanisms which may explain fatigue and others that may explain increases in "fitness". It is possible that cost of these adaptations may be globally quantified in the form of AE. This consideration of AE may help to better understand and develop new lines of inquiry around this topic of fatigue and adaptation. It is interesting that the GAS was applied to training independent from the concept of AE. However, we believe that this is likely due to the difficulty in defining what AE might be physically and how it would be measured or quantified. As Gorban et al. [12] have suggested, the desire to define AE as a physical quantity may impede the ability to consider its notion.

Within exercise science there is little mention of Goldstone's evolution of the concept of AE. Goldstone points out that:

"there is absolutely no shred of clinical evidence to suggest that about a year or more after a severe alarm reaction, a man has less power of adaptation to other stressors (such as accident, operation, or disease). If anything, the contrary appears to be true" [14].

We found Gorban et al.'s [11] discussion of Goldstone's axiom 1 enlightening and believe this model of adaptation might more appropriately describe what occurs in a human exercise model. Specifically, Goldstone suggests that AE can be created, but this ability will decrease with age. Perhaps a fitting example is skeletal muscle growth following exercise. An individual with no prior training experience may expect to see a large magnitude of skeletal muscle growth in a short period of time; however, this growth will eventually plateau. Counts et al. [15] suggests that the plateau in growth will likely occur within 3 months of training independent of the muscle group trained, and despite continued training efforts to grow this muscle. However, if an individual undergoes a period of detraining, the muscle will return to its baseline size [16]. If the skeletal muscle is then exposed to an additional training stimulus following detraining, this muscle will demonstrate a similar growth response as when it was previously untrained [16]. In addition, it has been demonstrated that with advanced age the muscle becomes anabolic resistant, meaning it is harder to induce skeletal muscle growth through the same training and nutrition. Thus, a similar stress will not result in a similar adaptation. There has been substantial molecular work to understand the mechanisms that drive growth and adaptation. However, exercise science should not discount the possibility of AE. Conversely, biological mechanisms of adaptation may help with the refinement and understanding of AE.

The application of the GAS in exercise science has limited itself to a physical interpretation of the experiments conducted by Selye. Gorban et al.'s work in this area [11,12] may help exercise scientists to consider signal transduction within an overall larger model of adaptation that includes AE as a governing principle as opposed to a physical quantity. This may allow for new advancements and innovation within the field of exercise science.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

None.

References

- [1] Cunanan Aaron J, DeWeese BH, Wagle John P, Carroll Kevin M, Sausaman Robert, Guy Hornsby W, et al. The general adaptation syndrome: a foundation for the concept of periodization. *Sports Med* 2018;48(4):787–97.
- [2] Selye H. Experimental evidence supporting the conception of “adaptation energy”. *Am J Physiol–Legacy Content* 1938;123:758–65.
- [3] Wilt F. Stress and training. *Track Tech* 1960;1:1–16.
- [4] Stone MH, O’Byrant H, Garhammer J, McMillan J, Rozenek R. A theoretical model of strength training. *J Strength Cond* 1982;4:36–9.
- [5] Wackerhage H. *Molecular exercise physiology: an introduction*. Routledge; 2014.
- [6] Buckner SL, Mouser JG, Dankel SJ, Jessee MB, Mattocks KT, Loenneke JP. The general adaptation syndrome: potential misapplications to resistance exercise. *J Sci Med Sport* 2017;20:1015–7.
- [7] Maclaren DP, Gibson H, Parry-Billings M, Edwards RH. A review of metabolic and physiological factors in fatigue. *Exerc Sport Sci Rev* 1989;17:29–66.
- [8] Drummond MJ, Fujita S, Abe T, Dreyer HC, Volpi E, Rasmussen BB. Human muscle gene expression following resistance exercise and blood flow restriction. *Med Sci Sports Exerc* 2008;40:691–8.
- [9] Phillips SM, Tipton KD, Aarsland A, Wolf SE, Wolfe RR. Mixed muscle protein synthesis and breakdown after resistance exercise in humans. *Am J Physiol* 1997;273:E99–107.
- [10] Vasenina E, Kataoka R, Buckner SL. Adaptation energy: experimental evidence and applications in exercise science. *J Trainology* 2020;9:66–70.
- [11] Gorban A, Tyukina T, Pokidysheva L, Smirnova E. Dynamic and thermodynamic models of adaptation. *Phys Life Rev* 2021;37:17–64. <https://doi.org/10.1016/j.plrev.2021.03.001>.
- [12] Gorban AN, Tyukina TA, Smirnova EV, Pokidysheva LI. Evolution of adaptation mechanisms: adaptation energy, stress, and oscillating death. *J Theor Biol* 2016;405:127–39.
- [13] Chiu LZ, Barnes JL. The fitness-fatigue model revisited: implications for planning short- and long-term training. *Strength Cond J* 2003;25:42–51.
- [14] Goldstone B. The general practitioner and the general adaptation syndrome. *S Afr Med J* 1952;26:88–92.
- [15] Counts BR, Buckner SL, Mouser JG, Dankel SJ, Jessee MB, Mattocks KT, et al. Muscle growth: to infinity and beyond? *Muscle Nerve* 2017;56:1022–30.
- [16] Psilander N, Eftestøl E, Cumming KT, Juvkam I, Ekblom MM, Sunding K, et al. Effects of training, detraining, and retraining on strength, hypertrophy, and myonuclear number in human skeletal muscle. *J Appl Physiol* 2019;126:1636–45.

Comment

Adaptation through the lens of single-cell multi-omics data

Comment on “Dynamic and thermodynamic models of adaptation” by A.N. Gorban et al.

Andrei Zinovyev ^{a,b,c,d,*}^a *Institut Curie, PSL Research University, F-75005 Paris, France*^b *INSERM, U900, F-75005 Paris, France*^c *CBIO-Centre for Computational Biology, Mines ParisTech, PSL Research University, 75006 Paris, France*^d *Laboratory of advanced methods for high-dimensional data analysis, Lobachevsky University, 603000 Nizhny Novgorod, Russia*

Received 20 April 2021; accepted 19 May 2021

Available online 21 May 2021

Communicated by L. Perlovsky

Keywords: Single-cell data; Omics data; Adaptation; Stress; Cancer

A.N. Gorban and his colleagues in their inspiring review described several theoretical models of adaptation and highlighted multiple examples convincing us in the existence of surprising at first thought phenomenon: the pattern of dynamical changes of basic statistical measures (correlation between features, their variance) can diagnose and prognose crises in the populations of objects exposed to stress [1]. The surprise is caused by the universality of this observation. Firstly, it is manifested in many different situations. Secondly and even more surprising, the features used do not have to be specifically designed to measure stress, even though the feature selection is still important. This suggests that the proposed models can serve as an insightful approach for Big Data analysis and interpretation.

One can note that all provided examples deal with macroscopic objects (people, patients, mice, plants, companies at the stock market). What if we will change the focus of the middle-out approach to a significantly smaller scale, e.g. from a tumor or a patient to a single cell? Would the suggested principles of adaptation thermodynamics still hold and what specific problems will arise?

Modern biotechnologies revolutionized biology and medicine by providing single-cell tools able to collect many thousands or even millions of quantitative traits per cell (typically, counts of RNAs transcribed from each gene in the genome, large-scale genomic modifications, profiling the chromatin state) [4]. Today we can take a piece of biological tissue (e.g., a tumor), dissociate it into millions of cells each of which is quantitatively characterized, for example, by expression (as counts of RNAs) of about 20000 distinct genes. Therefore, each biological sample can be represented as a multi-dimensional data point cloud in R^{20000} which properties can be characterized by gene-gene correlation graph and variance together with other characteristics.

DOI of original article: <https://doi.org/10.1016/j.plrev.2021.03.001>.

* Correspondence to: Institut Curie, INSERM U900, 26 rue d'Ulm, F-75248, Paris, France.

E-mail address: andrei.zinovyev@curie.fr.

Single-cell measurements are frequently used to study the effect of a stressful perturbation applied to cell populations. It can be a toxic treatment of cancer cells but also such general stresses as hypoxia or starvation. The stress can be extrinsic or intrinsic, e.g., resulting from non-physiological modifications of the cell genome. The main question here consists in how the cell population adapts to the caused stress by modifying the cell states and their distribution, and is the picture of high-dimensional omics measurements informative for characterizing the adaptation process.

The examples of such studies are numerous. We can briefly mention two recent ones from cancer biology. Early (in the first three days) adaptation of melanoma cell population was investigated after application of a potent targeted anticancer drug dabrafenib, belonging to the family of so-called BRAF inhibitors [2]. Most of the cells adapt to such treatment by entering quiescence but some “escapees” manage to maintain their cell cycle and eventually outproliferate other cells. Second example comes from Ewing sarcoma, a rare pediatric cancer, which is caused by a mutation translocating two chromosomes and leading to the appearance of a chimeric (non-physiological) transcription factor called EWS-FLI1. EWS-FLI1 can bind to short microsatellite sequences which are distributed more or less randomly in the human genome and perturbs the expression of thousands of genes. The effect of this perturbation is usually lethal for most cell types, but some (such as mesenchymal stem cells in bone marrow) can adapt, survive and even start to actively proliferate, giving rise to Ewing sarcoma cell populations. Experimentally we could modulate the expression of EWS-FLI1 and observe how the cancer cell population adapts [3].

It looks very tempting to apply the thermodynamic models described in [1], to study the adaptation in single cell populations. Indeed, similarly to human beings, individual cells can possess variable levels of adaptation energy and resources: therefore, most of the assumptions look valid. Moreover, the interplay between correlation and variance is commonly investigated in single-cell data analysis with such statistical concepts as overdispersion, differential correlation or differential variance being widely exploited [4]. Network entropy or dynamical covariation pattern analysis in genomic profiles have been used to anticipate crises that can manifest themselves in initiating, for example, carcinogenesis [5]. We showed that the local intrinsic dimensionality of data point clouds (measure related to the weight of the correlation graph) is an informative and biologically meaningful feature of single-cell datasets [6].

However, from our experience, interpreting the weight of the correlation graph and the variance changes computed from single-cell data, remains tricky. Let us just briefly describe some of the challenges that are faced in this way.

The first evident observation is that single-cell data are noisy and contain a large number of empty measures (so called technical drop-outs) [4]. These data are much noisier than the so-called “bulk” (using averages over millions of cells) measurements discussed in [1]. Various tricks such as using micro-pooling or variance stabilization are frequently used to improve the between-gene correlation structure in single-cell data, which, of course, impacts the quantification of the correlation graph.

Secondly, most omics-based single-cell measurements are destructive, which means that even if we track a cell population in time, we are not able to track the state of an individual cell. Here one deals with synchronic (as opposite to diachronic) data and usually not synchronized cell population snapshots. In single cell data science, a special family of computational methods, based on the concepts of pseudo-time, cell trajectories and data manifold alignment have been developed in order to match such snapshots from different timepoints and making them quasi-diachronic [4,7]. This way of treating the data might affect the conclusions of the analysis based on the weight of the correlation graph.

Thirdly and quite fundamentally, the role of selection and inheritance can be much more important when studying adaptation to stress at the level of cells rather than whole organisms. Thus, cancer cells in culture are rapidly dividing at the time scale of experimental perturbation, which is not the case in the examples described in [1]. The adaptive changes are inheritable to some extent after cell division. The difference between genetic and epigenetic inheritance conceptually consists only in the time scale for which the cell state modifications persist. Therefore, part of the changes in the structure of variance and correlations in a single cell population can be attributed to positive or negative selection. Whether to consider selection a part of adaptation, and also potential antagonism between instantaneous and population-based Darwinian fitness, need to be carefully considered in the analysis.

Forthly, defining cellular functional subsystems, an ingredient necessary for building thermodynamic adaptation models, can be not obvious. Methods based on matrix factorization and, in particular, Independent Component Analysis (ICA) seem to work well for splitting the set of genes into biologically meaningful groups with coordinated gene expression changes. In case of ICA, these collective changes are forced to be as statistically independent as possible [8], which reproducibly identifies such functional subsystems as DNA replication, mitotic programs (performing and exiting from mitosis), metabolic programs such as oxphos and glycogenesis, cell migration, response to hypoxia and others [3]. Scoring these functional subsystems defines a biologically tractable space of cell states, where one can ob-

serve dynamical coupling between functional subsystems under stress, increase of variance in cell state distribution, etc. Such representation of single-cell data might be more instructive than the use of the raw count data.

Finally, the most important in cancer biology cellular functional subsystem is the cell cycle machinery. Since the individual cells are usually not synchronized, the progression through the cell cycle creates a strong signal which can dominate the pattern of variance and correlation. In an ongoing study we observed that in slowly proliferating cell populations, the correlations between cell cycle genes as well as their total variance can be stronger compared to rapidly dividing cells. This is probably caused by that in rapidly dividing cells, some of the cell cycle-related transcripts do not possess enough time to be fully degraded after exiting from mitosis, which leads to their smaller expression amplitude. Interestingly, slowing down the cell cycle is a known cellular mechanism of adaptation to stress [9].

To conclude, we are convinced that the models of adaptation, based on the application of “middle-out” ideas suggested in [1] must become a working tool for making sense of single-cell omics datasets. However, further work, both theoretical and related to practical data analysis, is needed to take into account the specificity of the data and the studied systems, which should lead to verifiable and reproducible conclusions.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the French government under management of Agence Nationale de la Recherche as part of the “Investissements d’avenir” program, reference ANR-19-P3IA-0001 (PRAIRIE 3IA Institute), the Ministry of Science and Higher Education of the Russian Federation (Project No. 075-15-2020-808) and by European Union’s Horizon 2020 program (grant No. 826121, iPC project).

References

- [1] Gorban AN, Tyukina TA, Pokidysheva LI, Smirnova EI. Dynamic and thermodynamic models of adaptation. *Phys Life Rev* 2021;37:17–64. <https://doi.org/10.1016/j.plrev.2021.03.001>.
- [2] Yang C, Tian C, Hoffman TE, Jacobsen NK, Spencer SL. Melanoma subpopulations that rapidly escape MAPK pathway inhibition incur DNA damage and rely on stress signalling. *Nat Commun* 2021;12:1–14. <https://doi.org/10.1038/s41467-021-21549-x>.
- [3] Aynaud MM, Mirabeau O, Gruel N, Grossetete S, Boeva V, Durand S, et al. Transcriptional programs define intratumoral heterogeneity of Ewing sarcoma at single-cell resolution. *Cell Rep* 2020;30(e6):1767–79. <https://doi.org/10.1016/j.celrep.2020.01.049>.
- [4] Laehnemann D, Koster J, Szczurek E, McCarthy D, Hicks S, Robinson M, et al. Eleven grand challenges in single-cell data science. *Genome Biol* 2020;21(1):31. <https://doi.org/10.1186/s13059-020-1926-6>.
- [5] Teschendorff AE, Liu X, Caren H, Pollard SM, Beck S, Widschwendter M, et al. The dynamics of DNA methylation covariation patterns in carcinogenesis. *PLoS Comput Biol* 2014;10:e1003709. <https://doi.org/10.1371/journal.pcbi.1003709>.
- [6] Albergante L, Bac J, Zinovyev A. Estimating the effective dimension of large biological datasets using Fisher separability analysis. In: *Proceedings of the international joint conference on neural networks*. IEEE; 2019:N-19814.
- [7] Chen H, Albergante L, Hsu JY, Lareau CA, Lo Bosco G, Guan J, et al. Single-cell trajectories reconstruction exploration and mapping of omics data with STREAM. *Nat Commun* 2019;10(1):1903. <https://doi.org/10.1038/s41467-019-09670-4>.
- [8] Sompairac N, Nazarov PV, Czerwinska U, Cantini L, Biton A, Molkenov A, et al. Independent component analysis for unraveling the complexity of cancer omics datasets. *Int J Mol Sci* 2019;20(18):4414. <https://doi.org/10.3390/ijms20184414>.
- [9] Min M, Spencer SL. Spontaneously slow-cycling subpopulations of human cells originate from activation of stress-response pathways. *PLoS Biol* 2019;17:e3000178. <https://doi.org/10.1371/journal.pbio.3000178>.

Comment

Energy dissipation and storage in adaptation and homeostasis

Comment on “Dynamic and thermodynamic models of adaptation” by A.N. Gorban et al.

Armen E. Allahverdyan

Alikhanian National Laboratory (Yerevan Physics Institute), 2 Alikhanian Brothers street, Yerevan 0036, Armenia

Received 12 April 2021; accepted 18 May 2021

Available online 21 May 2021

Communicated by L. Perlovsky

Keywords: Adaptation; Homeostasis; Thermodynamic cost; Energy dissipation; Biological thermodynamics

The review by Gorban et al. [1] is unique in modern literature in that it summarizes the current knowledge on adaptation at four levels of description: (i) Observational and phenomenologic. (ii) Thermodynamic, which attempts at accounting for energy dissipation and entropy production during adaptation, and also proposes optimization principles for describing these phenomena. (iii) Axiomatic, which includes systems developed by Selye and Goldstone together with their subsequent ramifications. (iv) Kinetic, where models inspired by chemical kinetics are introduced and studied for revealing the real-time dynamics of adaptation.

This comment has two related goals. First, I would like to make several clarifications on the notion of adaptation energy reviewed in [1]. They relate to the difference between homeostasis and adaptation. I shall also emphasize open issues in (ii), which necessitates new models and approaches from both physical and biological viewpoints. Second, I want to outline relations between axioms by Selye and Goldstone and a related axiomatic system proposed in 1930's by Bauer [2], and developed since then by biophysicists and physiologists [3,4,7,5,6]. Certain ideas which are common to both axiomatic approaches were worked out in more detail in the second school of thought. Hence their joint development might be a topic of future work.

Let me start with two definitions; see e.g. [8]. A state is called homeostatic, if it is stabilized against a class of perturbations of extensive (dynamic) variables, e.g. concentrations of chemical elements or configurational degrees of freedom. Likewise, a state is called adaptive, if it is stabilized against perturbations of intensive variables, e.g. temperature or a chemical potential. Such variables frequently appear via parametric, environmental constraints, hence adaptation deals with environmental perturbations. These working definitions are not exhaustive, at least because both homeostasis and adaptation should refer to specific functions (e.g. energy transfer or storage), which are going to be stable. But they roughly suffice for our purposes.

DOI of original article: <https://doi.org/10.1016/j.plrev.2021.03.001>.

E-mail address: armen.allahverdyan@gmail.com.

Now the difference between homeostasis and adaptation is that for the former it is possible to envisage fully equilibrium mechanisms, while for the latter it is impossible [9–12]. This difference is deeply rooted in thermodynamics and has to do with the Le Chatelier-Braun principle [10]. In contrast, to what many people think and what is usually presented in textbooks [13], this principle is not just a statement on specific stability inherent in the thermodynamic equilibrium. The principle first of all divides between stability with respect to extensive *versus* intensive variables [14]. For the first case, Le Chatelier-Braun principle is indeed a statement on stability, but for the second case the principle states *thermodynamic instability* [10,14]. This difference implies that while there can be passive (equilibrium) mechanisms for homeostasis, there cannot be equilibrium mechanisms for adaptation, since the latter deals with perturbations of environmental (intensive) parameters such as temperature and/or chemical potential [10]. A non-equilibrium situation is necessary for neutralizing the Le Chatelier-Braun principle and achieving adaptation [9–12]. One such scenario of adaptation was explored in [10,12], and it relates to the energy stored in non-equilibrium control degrees of freedom: this stored energy prevents the validity of the principle and ensures a stability with respect to environmental perturbations (i.e. adaptation).

As already emphasized in [1], this scenario of adaptation closely relates to the concept of adaptation energy proposed by Selye. In this axiomatic system the energy is a resource that is consumed during the adaptation process [1]. A similar, but stronger condition holds for models studied in [9–12]: the energy is not only dissipated in the course of adaptation (i.e. transiently), but also permanently, i.e. in the adapted steady state [9–12]. Such a high energy cost is a feature of current stochastic models of adaptation. I propose that it is not realistic, because it puts too much burden on the living organism, though sufficiently general estimated of the adaptation energy cost on the level of a single organism are currently lacking [1].

The same question on the minimal amount of energy dissipation can be asked about active (cyclical motion-driven) scenarios of homeostasis. Here however we can be more precise in posing the energy cost problem: once scenarios of active homeostasis involve the very process of (useful) energy storage, it is highly unlikely that such processes will be accompanied by permanent dissipation of energy.

Thus the question on the real amount of energy costs in adaptation and active homeostasis is open. Without clarifying these questions, it will be impossible to establish a real correspondence between axiomatic structures of adaptation and microscopic models. It should also be noted that the above link with the Le Chatelier Braun principle is yet model-dependent [10]. Its sufficiently general formulation is yet to be achieved.

Yet another pertinent open problem relates to the origin of the energy employed for adaptation. A usual reference to sources of non-equilibrium energy in cell biophysics (ATP) is certainly not sufficient here, e.g. because it is known that the energy needed for adaptation is not easy to recover [1]. Selye even thought that it is not recoverable at all. What is the physics of devices that extract energy from non-equilibrium resources and employs them for their own adaptation? How large such devices have to be?

In the last part of this comment I would like to mention that in 1930's Bauer proposed an axiomatic system summarized as (1)–(4) below, and employed it for various problems in theoretical biology [2]; see [3] for a recent review. This system does resemble the those by Selye and Goldstone [see [1] for review], but is potentially richer.

(1) A living organism contains a reservoir of stored energy that is employed for maintaining its own non-equilibrium (living) state and for doing external work. The stored energy is not the usual free energy, since the latter is present in equilibrium as well. From the viewpoint of a modern thermodynamics, the notion of the stored energy resembles the energy kept at certain negative temperature, because it is capable of doing work in a cyclic process [15].

(2) Living organisms do not directly employ the chemical energy of food for doing work. Instead, the chemical energy is used for rebuilding the stored energy reservoir. Bauer motivates this axiom from Rubner's result (calorie is a calorie): the heat production by the organism in its natural conditions is the same as it would be upon burning this same food [2]. This axiom was proposed even before the concept of ATP was discovered, and it is consistent with modern bioenergetics [4].

(3) The stored energy per mass unit (i.e. the stored energy density) is maximal at birth, and then it normally decays monotonically. However, during the organism growth the overall stored energy increases. Once the growth is over, both the stored energy density and the overall stored energy decrease with a speed determined by an organism itself. This is the aging. Thus Bauer proposes that there is a maximal amount of food energy (per mass unit) each organism is able to process [2]. According to modern estimates, for mammals this energy is about 191600 kcal per 1 kg of weight [5,6]. Humans do not fit to this estimate, since for them this number is some 4 times larger [6].

(4) Under certain specific conditions there are processes that prevent aging. In such processes the overall mass of the organism can shrink thereby increasing the stored energy density. This concentration process is seen e.g. in fasting. This axiom found applications in gerontology [7], though its integration with current aging theories [16] is not yet achieved.

Yet another exclusion from (3) was described by Arshavsky in the development of skeletal muscles [5,6]; he pointed out that due to muscle activity *both* the stored energy density and the overall stored energy of the muscle can increase. This effect is visible for skeletal muscles only under moderate (not excessive) external stress [5,6]. However, there is however an uncertain point here that needs further attention: Refs. [5,6] seem to mix the notion of stored energy with that of free energy. It is thus highly desirable to reconsider statements by Arshavsky via the modern knowledge on energetics of skeletal muscles; see e.g. [17].

To summarize this comment, Ref. [1] motivates the following research program towards the physics of adaptation: (1) A more precise understanding—via the Le Chatelier Braun principle [14] or its relatives [18,19]—for the necessity and the amount of energy costs in adaptation as compared to homeostasis. (2) Development of more general axiomatic schemes for adaptation energy and its change integrating approaches by Selye and Goldstone with those of Bauer. Such integrated schemes can motivate more refined and interesting models.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the SCS of Armenia, grant No. 20TTAT-QTa003.

References

- [1] Gorban AN, Tyukina TA, Pokidysheva LI, Smirnova EI. Dynamic and thermodynamic models of adaptation. *Phys Life Rev* 2021;37:17–64. <https://doi.org/10.1016/j.plrev.2021.03.001>.
- [2] Bauer ES. Theoretical biology. Moscow: VIEM; 1935.
- [3] Grandpierre A, Chopra D, Kafatos MC. The universal principle of biology: determinism, quantum physics and spontaneity. *Neuroquantology* 2014;12(3).
- [4] Blumenfeld LA, Tikhonov AN. Biophysical thermodynamics of intracellular processes: molecular machines of the living cell. Springer Science & Business Media; 2012.
- [5] Arshavsky IA. Focus for development: the energy rule of skeletal muscles. *Dev Psychobiol* 1974;7:291–5. <https://doi.org/10.1002/dev.420070403>.
- [6] Arshavsky IA. Theory of individual development (biophysical aspects). *Biophysics* 1991;36(5):872–84.
- [7] Voeikov VL. Bio-physico-chemical aspects of aging and longevity. *Adv Gerontol (Uspekhi gerontologii)* 2002 Jan 1;9:54–66.
- [8] Viktorov VA, Novoseltsev VN, Shumakov VI. Control in and for biosystems. *IFAC Proc Vol* 1984 Jul 1;17(2):2985–3002.
- [9] Lan G, Sartori P, Neumann S, Sourjik V, Tu Y. The energy-speed-accuracy trade-off in sensory adaptation. *Nat Phys* 2012 May;8(5):422–8.
- [10] Allahverdyan AE, Wang QA. Adaptive machine and its thermodynamic costs. *Phys Rev E* 2013;87:032139.
- [11] Sartori P, Tu Y. Free energy cost of reducing noise while maintaining a high sensitivity. *Phys Rev Lett* 2015 Sep 8;115(11):118102.
- [12] Allahverdyan AE, Babajanyan SG, Martirosyan NH, Melkikh AV. Adaptive heat engine. *Phys Rev Lett* 2016;117:030601.
- [13] Landau LD, Lifshitz EM. Statistical physics, part 1. Pergamon Press; 1980.
- [14] Gilmore R. Le Chatelier reciprocal relations and the mechanical analog. *Am J Phys* 1983 Aug;51(8):733–43.
- [15] Allahverdyan AE, Balian R, Nieuwenhuizen TM. Maximal work extraction from finite quantum systems. *Europhys Lett* 2004 Aug;67(4):565.
- [16] Yates FE. Theories of aging: biological. In: Birren JE, editor. *Encyclopedia of aging*. UK: Elsevier Inc.; 2007. p. 601–10.
- [17] Sugi H, Pollack GH, editors. Mechanisms of work production and work absorption in muscle. Springer Science & Business Media; 2012.
- [18] Gorban AN, Gorban PA, Judge G. Entropy: the Markov ordering approach. *Entropy* 2010 May;12(5):1145–93.
- [19] Allahverdyan AE, Galstyan A. Le Chatelier's principle in replicator dynamics. *Phys Rev E* 2011 Oct 13;84(4):041117.



Reply to comment

It is useful to analyze correlation graphs

Reply to comments on “Dynamic and thermodynamic models of adaptation”

A.N. Gorban ^{a,b,*}, T.A. Tyukina ^{a,b}, L.I. Pokidysheva ^c, E.V. Smirnova ^c^a Department of Mathematics, University of Leicester, Leicester, UK^b Lobachevsky University, Nizhni Novgorod, Russia^c Siberian Federal University, Krasnoyarsk, Russia

Received 13 October 2021; accepted 25 October 2021

Communicated by L. Perlovsky

0. Introduction

What can be considered as the main result of a long research program? Let us take a look at great examples. According to Arnold [1], Newton's fundamental discovery, the one which he considered necessary to keep secret and published only in the form of an anagram, in contemporary mathematical language means: “It is useful to solve differential equations”. The discovery and elaboration of a new tool, with a demonstration of its wide applicability, is a good main result when supplemented by many smaller advances that demonstrate the usefulness of the developed tool and should be important in their own right.

In 1987, we analyzed the changes in correlation graphs between various features of the organism during stress and adaptation [2]. Results of 33 years of work of many researchers and groups were summarized in the review paper [3]. In this review, we aimed to demonstrate that it is useful to analyze correlation graphs. In addition, we should add that the concept of adaptability (‘adaptation energy’) introduced by Selye in 1938 [4] is useful, especially if it is supplemented by recently invented ‘adaptation entropy’ and free energy, as well as by an analysis of limiting factors.

However, the proper extended review and outlook were finalized when the twelve additional comment papers were published together with our article [5–16]. Our ‘reply to comments’ is an introduction to these twelve supplementary materials. There are several aspects in which commentators have enriched our review:

DOI of original article: <https://doi.org/10.1016/j.plrev.2021.03.001>.

DOIs of comments: <https://doi.org/10.1016/j.plrev.2021.05.009>, <https://doi.org/10.1016/j.plrev.2021.04.003>, <https://doi.org/10.1016/j.plrev.2021.04.001>, <https://doi.org/10.1016/j.plrev.2021.05.005>, <https://doi.org/10.1016/j.plrev.2021.05.002>, <https://doi.org/10.1016/j.plrev.2021.05.004>, <https://doi.org/10.1016/j.plrev.2021.04.005>, <https://doi.org/10.1016/j.plrev.2021.05.008>, <https://doi.org/10.1016/j.plrev.2021.04.004>, <https://doi.org/10.1016/j.plrev.2021.07.001>, <https://doi.org/10.1016/j.plrev.2021.05.010>, <https://doi.org/10.1016/j.plrev.2021.04.002>.

* Corresponding author at: Department of Mathematics, University of Leicester, Leicester, UK.

E-mail addresses: a.n.gorban@le.ac.uk (A.N. Gorban), tt51@le.ac.uk (T.A. Tyukina), pok50gm@gmail.com (L.I. Pokidysheva), seleval2008@yandex.ru (E.V. Smirnova).

<https://doi.org/10.1016/j.plrev.2021.10.002>

1571-0645/© 2021 Elsevier B.V. All rights reserved.

- Presentations of materials on changes in correlation networks under stress, illness and adaptation, which analyze various situations and use additional tools for data preprocessing and analysis of correlation graphs. These extensions are important for proof that analysis of correlation graphs is useful in studying of adaptation.
- Suggestions to extend methods onto different phenomena and problems, from aging and single cell omics to molecular networks and other intriguing systems.
- Discussion of alternative or deeper backgrounds of the effects described in our review, and description of the conceptual problems that the proposed methodology will face in the further development and expansion of application areas.

Of course, each comment contributed to several aspects of the content. Below, we briefly discuss these enrichments.

1. Additional materials on the main area of the review

Dynamical network biomarkers A treasure box of “dynamical network biomarkers” was opened for readers in [14]. This concept was developed to characterize the critical state during the disease progression. We agree that these ideas are consistent with our correlation graph approach. Moreover, taking into account the publication dates, we are proud to mention that the dynamical network biomarkers present the further development of our approach. Our series of publications started in 1987 (see, for example, [2,17–22]). The approach was later applied to many systems and elaborated in various directions by many authors, with references to our work [23–35] or with rediscovery of some ideas [36–38] (for the more detailed review we refer to [3]). Of course, each paper was not just an application or rediscovery of previously known ideas. Most groups of authors proposed their own versions of methods suitable to their material and developed some theoretical backgrounds for their findings.

In particular, the important new tools for selection of the relevant dominant or leading subnetworks in the large networks of attributes near crisis were developed by comparison between the internal (within the group) and external correlations. A brief review of the theory and practice of dynamical network biomarkers was presented in the comment [14]. A different approach for selection of the most relevant subnetworks of attributes is based on principal component analysis [26,27]. In this approach, at the first stage, the most important principal components for disease recognition should be found (the components with the greatest significance for the classification task), and then the attributes that dominate in these principal components should be selected for further correlation and variance analysis (for more details, we refer to the original works).

The series of work presented in [14] systematically uses the concept of attractor bifurcations as a hidden cause of the observed transformation of the network. The general theory of bifurcation of attractors and critical delays in general dynamic systems (with and without noise) was developed in 1978-1980 in the PhD thesis of one of the authors [39,40]. (An abridged version of the thesis [39] was published later in English [41], see also a brief review [42] with further references.) However, we found that a systematic theory of the transformation of attribute networks under adaptation and stress can benefit from the concepts developed by Selye and his school. Below, we discuss this question and the non-trivial relations between attractor bifurcations and the concept of adaptability in more detail using also other comments that considered bifurcations of attractors as one of the main mechanisms of the observed effects [6].

Someone’s noise is another one’s signal K. Selvarajoo in his comment [12] discussed high-throughput gene expression data analyzed to interpret and predict the complex immune, cancer and biofilm responses. In general, he supported our findings but attracted more attention to biological noise. This noise can be independent from the technical noise like instrumental effects and can be a signal of approaching critical transitions. We completely agree with this comment. Moreover, the correlation graph approach discussed in our review with analysis of correlations and variances is a tool for study of sensible noise in the networks of attributes. These methods are widely used and modified for different purposes. For example, correlations and variances both in the organism-to-organism [2,19,29] (or cell-to-cell [12]) variations and in the fluctuations of the attributes of one organism in time [32] can provide important information about adaptation and crisis. The further development of this methodology provides inspiring examples even from applied humanities, such as the analysis of fluctuations in time of physiological parameters (cardiac activity and skin conductance) in a ‘psychotherapist-patient’ pair, which measures the level of empathy and may predict the

success of psychotherapy [43], or the analysis of correlations between variations of public fears for different regions of the country that can predict political crises [34].

This topic, analysis of noise and individual behavior of various structural components (both material, like cells, tissues, etc., and virtual, like principal components, biomarkers, and subnetworks) was stressed by many commentators. C. Franceschi stated that “*each cell has a specific signature and present subtle differences with respect to the other cells derived from the same source,*” – and these differences increase with age [5]. In this comment, like in [12], our approach is friendly criticized for insufficient attention to this heterogeneity of life system: “A major usually non explicit assumption of the approach adopted by Gorban et al. is that the elements of each specific cellular network are basically similar each other’s. This basic assumption is likely not true.” We fully accept the fundamental heterogeneity of many life systems at all levels but cannot accept that their homogeneity is our ‘major assumption’.

However, the source of the impression that our systems are “basically similar” is very clear (to us): To apply the correlation graph approach to a group of biological objects (organisms, cells, etc.), no one needs their homogeneity. In fact, ‘*meta-uniformity*’ is desirable instead: *the attributes selected for analysis must be measurable for all objects*. Of course, some palliative solutions can be used when attributes are immeasurable for some objects, for example, a special value can be entered meaning ‘impossibility of measurement’. However, the correlation of the two attributes can only be calculated for groups (or subgroups) where both attributes are measurable. Technically, one can consider a ‘crazy kilt’ of such subgroups with a different set of objects for different pairs of attributes, but meta-homogeneity with the same attributes for all objects makes data analysis much easier and more transparent.

According to the comment [6], the crucial importance of fluctuations in living systems is caused by the “small integer numbers of key and lock molecules” of many types. “The key issue is that randomness is a central feature of the microscopic level, and it is therefore unlikely that microscopic details could help in understanding the logic behind living organisms.” Therefore, the special mesoscale approach is needed, that does not rely on the central limit theorem and mass action law but, on the other hand, is not as noisy as the microscopic level is [6]. Despite of that, another comment [10] is focused on the single cell level and discussed the data collections that are produced at that ‘microscopic’ level by single cell omics as a promising area of applications of the correlation graph approaches and their further development. The comment [8] goes even deeper, to more ‘microscopic’ and noisy molecular networks. The fluctuating world of many personalized objects can be represented as a supernetwork, that is a network of networks. Promising approaches for constructing such supernetworks of potentially infinitely depth hierarchy were discussed in the comment [7]. This hierarchy starts from the graphs of binary correlations (or their close analogues) and then goes deeper to more complex structures.

Separation of real life high-dimensional and heterogeneous data into signal and noise is extremely problematic and can be never done for sure. The important data mining ‘axiom’, *Someone’s noise is another one’s signal*, should be included into undergraduate textbooks, and biological networks can provide good examples of transforming apparent noise into important signal.

2. Extending methods to new phenomena and problems

Dynamics of aging According to the comment of C. Franceschi, [5], the approach presented in our work “will help not only in disentangling the complexity of aging (particularly in humans) but also in establishing a new ‘aging medicine’, and above all in identifying new tools capable of early predicting power regarding human health trajectories.” In this comment, the phenomena of aging that can be covered by our approach were analyzed and some open problems were formulated.

The concept of adaptation energy and optimization of its spending discussed in the paper “fits very well with what I learned from the study of human aging and longevity,” wrote one of the world’s leading experts on aging. This opinion supports our claim that the concept of adaptability (‘adaptation energy’) is useful. It could be worth to mention that a very recent critical review of adaptation energy and its applications to exercise science [44] states that this concept is applicable to the phenomenon of aging rather than to sports.

C. Franceschi drafted a research program for application of dynamic and thermodynamic models of adaptation to aging phenomena. As per his comment, the main and urgent challenge to our approach, is “the fundamental heterogeneity and uniqueness of humans at all biological level (genetic, epigenetic, metabolic, among others) which emerges whenever the results of a study are analyzed at the individual level,” and an enormous amount of work is expected to meet this challenge. We have already begun work on the analysis of large heterogeneous systems and have cre-

ated open access software tools for identifying development trajectories in big data [45–48]. These tools can help in predicting “human health trajectories”, as requested by the commentator. Nevertheless, there remain more inspiring questions than answers in the theory and practice of optimal aging and longevity [5].

Exercises and sport Exercise and sports data provide a wealth of material for the study of adaptation. Any theory of adaptation is incomplete without assimilation of these materials. We were deeply concerned that Selye’s classical concept of adaptation energy did not seem to be able to describe the training phenomena that are well known in the science of exercise and beyond to anyone with a sporting background. Similar concern was formulated by several authors, see a review [44]. Our commentators [9] are experts in exercise science. They analyzed the concepts of adaptation from the “exercise science perspective”. We are grateful for an accurate understanding of our modification of the concept of energy of adaptation and for examples that demonstrate the need for modification. Here, we would like to add one more comment. To capture the apparent training phenomena we had to introduce “free energy of adaptation”. It included two components: adaptation energy and adaptation entropy. The entropy of adaptation is necessary for the formalization of adaptive changes in the scale of stressors. After training, the organism can withstand larger loads. The same load can be ‘small’ for a trained body and ‘large’ (or even impossible) for an untrained one. This rescaling of the stressors can be measured by a logarithm of the rescaling coefficient (logarithm is needed for additivity). This is a typical entropic measure – logarithm of available ‘volume’. The dynamics of the load scale described by our free energy models [3] are qualitatively consistent with the phenomena discussed in the comment [9], but additional efforts are required for detailed modeling and analysis to verify the quantitative agreement.

Single cell omics We dealt [3] with whole organisms (mice, people, plants) or “super-organismic” objects (stock market or social systems). A. Zinovyev proposed to change the focus to a significantly smaller scale, to a single cell [10]. This is not just an interesting and important question asked by many commentators. The comment [10] proposed to work with single-cell omics. This is a new technology that gives information about genomes, gene expression, and cell heterogeneity for thousands of individual cells from a single organism, simultaneously. This revolutionary technology brings a huge amount of new data we have never seen before. It is a challenge to biology and medicine to assimilate these datasets. Several specific difficulties of single cell omics data analysis were discussed in the comment [10]. In particular, individual cells are usually destroyed when measured, so we can track the population of cells over time “on average”, but the state of an individual cell can only be observed once. It is very desirable to join individual cells (represented by points in a multidimensional space) into trajectories of their development. These trajectories can branch and the whole picture is expected to be a bunch of branching cell development and differentiation trajectories. Extraction of this bunch from a single cell data cloud is a special data mining problem and several tools are developed (for example, the method of topological grammars and principal graphs [45,49]). However, this is not the end of the story. The motion of cells in the feature space may include also oscillation in the cell cycle that is different from the development and differentiation. This is especially important for growing cell populations like cancer. Special techniques are needed to separate these two types of motion: cell cycle and development with differentiation. In a recent work [50] a motion separation approach based on independent component analysis was developed. Universal mathematical [51] and computational [52] models of cell cycle are useful for such separation. The developed methods for processing of single cell omics data open the door for detailed analysis of adaptation and stress at the single cell level. A research program of this analysis is drafted in the comment [10] and we will be happy to see the results and participate in this new fascinating intellectual travel.

Molecular networks Can we go deeper than the single cell level? The comment [8] demonstrates one more step: from cells to molecular networks. Adaptation at the level of protein networks is possible. Recent works reveal that protein networks adapt to mutations. They develop new allosteric networks. For us, this is a completely new angle of view, and we are more than happy that the integrated dynamic and topological approach works at that level and helps in analysis of adaptation at the level of protein networks. As a result, we can imagine the future construction of adaptation science, from the very bottom level (molecular networks) to the single cell adaptation, then the tissue and organismic levels, and finally, the level of super-organismic entities. In this context, the concept of “network of networks” [7,53] seems to be very useful.

3. Alternative or deeper backgrounds of the observed effects

Bifurcations versus adaptability dynamics In two comments, [6,14] the bifurcations of attractors are considered as the main mechanism of the transformation of correlation graphs near the ‘tipping points’. In our review [3], we discussed three proposed mechanisms of the observed effects: dynamics of adaptability described by the factors–resource models, approaching bifurcations of hidden dynamic systems, and approaching the border of survival (in the framework of Fokker–Planck models). All these mechanisms describe some experimental data, at least, qualitatively. Moreover, even if some data are not described, there remains a hope the model can be modified accordingly to meet better the real phenomena.

It is difficult to select the best type of models because all the models refer to hidden phenomena that are not observable directly. For example, the adaptation energy is not a physical quantity and cannot be measured directly. Instead, adaptation energy (or free energy) should be considered as an internal coordinate on the *dominant path* in the model of adaptation [22]. Many questions can be asked: What is the dynamical system that has this dominant path, what is its phase space, why the dominant path appears? (This can be the result of either averaging, or the existence of a slow manifold, or a combination of these reasons.)

Very similar questions can be asked for the bifurcation model. Nobody has observed the underlying dynamical systems and its ‘network of attractors’. Therefore, all this reasoning remains a useful metaphor. The situation is very similar to the ‘self-organized criticality’ (SOC) that provides us with a metaphor and simple nice examples like a sandpile but does not create quantitative models. “One of the main misconceptions about SOC is to think that it is a traditional quantitative theory and that, as such, it can yield detailed quantitative predictions once confirmed that the system exhibits SOC dynamics [54].”

Existence of the dominant path and a single parameter that describes behavior of a complex system is usual near bifurcation and phase transitions. Perhaps, the first famous example gave Lev Landau in his work about stability loss of laminar flows [55]. In this work, he found the equation for the deviations from the steady laminar flow without explicit use of the Navier–Stokes equations. The closeness to the bifurcation was the main assumption. In that sense, two approaches, bifurcation and a special dominant path with adaptability as a main coordinate, are not in principal contradictions. Fluctuations and critical delays become very important near bifurcations, here we agree with our commentators. (See also our detailed analysis of attractor bifurcations and critical delays in dynamical systems, [39–42]. In particular, we analyzed the effect of fluctuations modeled by ε -trajectories.)

Despite of some similarity, there exists a difference between just a general ‘closeness to a bifurcation’ and a specific dominant path factor-resource models. Instead of rephrasing our large paper [3], we will present a very recent simple example of a social system, where these factor-resource relations are obvious.

Development of world-class universities in China was analyzed in [56]. This process was started in 1998 and recently, high positions of Chinese universities in world ranking are widely known. At the beginning of this race for ranking, vast investments of money have indeed earned China’s universities ‘quick wins’. In that sense, one factor dominated the accelerated development of the universities to the relatively high level. But after that, more individual institutional and cultural factors determine the development of the universities, “including dominion over university-level strategic initiatives, promotion, tenure, research directions, establishment of new academic programs, and publication and academic content” [56]. Thus, efforts and successes of different universities become different and the race for higher success is not dominated just by funding. According to [56], the simple government’s command-and-control approach to higher education institutes became largely inconsistent with the diverse world of universities and at this “multifactorial” stage, a more flexible approach is necessary. We cannot comment on these advices about management, but the difference between the single-factor development and polyphonic multifactorial motion seems to be obvious.

The comments [6,14] include much more material. In particular, [6] presents a very dense review of many mathematical, biological and even epistemological problems in understanding biological adaptation with special attention to stochastic processes and limited predictive power of many classical families of models like ordinary differential equations of chemical kinetics. (The comment [12] also stressed that these models are insufficient in the stochastic world of genome activity.) High importance of fluctuations was demonstrated also in the comment [14]. All these problems are very important, indeed. Here we would like just to stress that relatively simple ‘middle level’ models may work well despite the complexity and stochasticity of the low-level detailed picture.

Statistical physics of life and network thermodynamics The factor-recourse models work like thermodynamics, whereas the hidden bifurcation of unknown attractors is a metaphor of statistical physics. A. Giuliani in his comment [13] presented an idea of ‘organized complexity’ and ‘statistical physics of life’. He proposed that the foundation of a biological statistical mechanics “should focus on the changes in correlation patterns marking state transitions”. At the same time, the network thermodynamics of life is necessary to consider changes in macroparameters. For this purpose, we utilized the typology of systems of factors and their interactions proposed by Tilman [57]. We developed these ideas further and applied to analysis of adaptation. In addition to the Liebig’s law of the minimum, various different interactions of factors are possible like synergy. The rich family of “factor-resource” models can work not only as a metaphor, but also provide quantitative models of real systems. To overcome the quasistatic nature of these models (‘periodic adaptation’), a family of dynamic models of individual adaptation was proposed.

The manifesto of statistical physics of life [13] supports the main results: It is useful to analyze correlation graphs and thermodynamic models based on adaptability (‘adaptation energy’) concept. It also opens further methodological perspectives.

Networks of networks The concept of “networks of networks” [7] covers many ideas expressed by the commentators. The hierarchy of networks that begins with molecular networks [8] at the lowest level, then continues with networks of cells and their features [10], then with correlation networks of the attributes of an organism [2,3,14,19], and so on, presents an example of such a network of networks. Analyzing these constructs may require special mathematical and statistical tools.

In our paper [3], we considered the correlation graphs, where the nodes were attributes and edges represented the correlations between the attributes. The simplest definition of edge weight is Pearson’s correlation coefficient. In our previous works, we tried many various characteristics and, finally, were satisfied by the simplest version. Nevertheless, this is not a compulsory choice and for formal description of ‘networks of networks’ various different approaches may be proposed. Informational characteristics, like information gain, may be useful, especially for qualitative attributes. Various regression residuals are used in the “parencletic” (from the Greek for “deviation”) and synolytic (from the Greek for “ensemble”) network representations [7]. These networks were successfully used in medical applications. Essentially, they are also the correlation networks but with the different measures of correlations. However, the details matter and different measures may be useful in different situations. It will be important to compare different approaches and we hope that the comment [7] could be extended to an analytical work about different versions of correlation graphs and their performance analysis. Even more, the ensemble of different correlation graphs can be used for description of a complex biological system, and we may not need to choose the best correlation graph, but rather use all of them together [7].

Another technique is discussed in the comment [15]. Many modern tools are developed for learning of intellectual agents, including neural networks. Adaptation can be modeled by learning. Evolution of a population of such agents with small mutations of ‘genotypes’ and optimization of phenotypes by learning can clarify the problem of instantaneous fitness discussed in our work. These systems of intellectual agents can be used for modeling networks of networks.

Three (or four?) faces of free energy in adaptation and stress Two free energies were used in our work. One of them is the classical physico-chemical free energy used in [3, Appendix A] for demonstration of ‘thermodynamic cost of homeostasis’. Another is the adaptation free energy, combined from the adaptation energy and the adaptation entropy introduced for modeling of the adaptive rescaling of stressors in training. Two comments, [11,16] added a lot to this topic.

The main challenge, formulated in [11] is “to establish a real correspondence between axiomatic structures of adaptation and microscopic models.” According to this comment, the difference between homeostasis and adaptation is the difference between two types of control: for homeostasis, this is control of dynamic variables (‘thermodynamic coordinates’), like concentrations, and for adaptation this is control of the conjugated variables (‘thermodynamic forces’), like chemical potentials. Behind this definition, we found the representation of biological systems as *controlled dynamical system of non-equilibrium thermodynamics*. We fully support this basic idea. Adaptation in dynamical systems [58] is important for analysis of biological processes, and these dynamical systems must obey thermodynamics. However, realization of this idea still requires a lot of work and further details, even in the basic definitions.

The comment [11] presents Bauer's axiomatics of adaptation [59] and explains, why it meets the physico-chemical backgrounds of biological adaptation better than Selye (or Selye–Goldstone) axioms. A research program “towards the physics of adaptation” is proposed that consists of two items. We reformulate them as follows:

1. Develop accurate thermodynamic estimates of energy consumption in adaptation and homeostasis (according to [11], the existing models overestimate this consumption, especially for adapted steady states).
2. Integrate the Selye–Goldstone phenomenology with the physico-chemical approach to adaptation proposed by Bauer (and the modern versions of these approaches) and create more realistic and interesting models.

It seems that the comment [16] approaches thermodynamics of adaptation from the opposite side. The concept of variational free energy is clearly explained from scratch, from variational Bayesian inference. The formal definition [16, Box 1] is followed by an explanation of qualitative details, meaningful analogies and even metaphors. According to [16], the variational free energy evaluates the surprisal under optimism-biased expectations from the world, that is, it measures instantaneous distress. It can be used for deriving the instantaneous fitness using the learning metaphor of adaptation (compare to the remark from [15] about evolutionary learning of intellectual agents and models of adaptation and evolution). It is impressive that the dynamics of “Bayesian thermostats” in active inference demonstrate the behavior of variance and correlations between intellectual agents that was observed in many living systems under stress in adaptation process.

Finally, the concept of variational free energy meets physical thermodynamics in the following form: in practice of modeling living systems, the variational free energy and the stored thermodynamic free energy have *inverse relationships* because “the minimisation of variational free energy in living agents is associated with the build-up and maintenance of the physical structure realizing these agents.”

The combination of comments [11,16] gives hope that a new revolutionary concept of variational thermodynamics of adaptation will soon appear, taking into account both non-equilibrium thermodynamics and Bayesian learning.

A very minor answer to a minor comment from [16] is ethically important for us. In a brief discussion of our estimate of free energy cost of a non-equilibrium state [3, Appendix A], an earlier work of Horowitz et al. (2017) [60] was mentioned in [16]. We were ready to cite our book (1984), which contains these estimates (and much more) [61] (some results of this book were later published in English [62]), but we were ashamed when found that the first results of this type were published in 1935 by Bauer [59] (not cited in our review). We must never forget this very talented researcher who was arrested and killed on the peak of his creative life.

4. Instead of conclusion

After 33 years of research of many authors, discoveries and rediscoveries, we can say with complete confidence: It is useful to analyze correlation graphs. The combination of comments creates a polyphonic effect around a very deep problem of adaptation. In this ‘reply’ we tried to share our impressions and enjoyment of this polyphony. We are very grateful to all commentators for new questions, tools, hypotheses, and answers to some problems. The main expected result of our joint efforts will be “the enlargement of the spectrum of the phenomena amenable of a quantitative approach” [13].

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The work was supported by the Ministry of Science and Higher Education of the Russian Federation (Project No. 075-15-2021-634).

References

- [1] Arnold VI. Geometrical methods in the theory of ordinary differential equations. New York: Springer; 2012.

- [2] Gorban AN, Manchuk VT, Petushkova (Smirnova) EV. Dynamics of physiological parameters correlations and the ecological-evolutionary principle of polyfactoriality. *Problemy ekologicheskogo monitoringa i modelirovaniya ekosistem* (The problems of ecological monitoring and ecosystem modelling), vol. 10. Leningrad: Gidrometeoizdat; 1987. p. 187–98.
- [3] Gorban AN, Tyukina TA, Pokidysheva LI, Smirnova EV. Dynamic and thermodynamic models of adaptation. *Phys Life Rev* 2021;37:17–64. <https://doi.org/10.1016/j.plrev.2021.03.001>.
- [4] Selye H. Adaptation energy. *Nature* 1938;141(3577):926. <https://doi.org/10.1038/141926a0>.
- [5] Franceschi C. Aging, inflammaging and adaptation: comment on “Dynamic and thermodynamic models of adaptation” by A.N. Gorban et al. *Phys Life Rev* 2021;38:107–10. <https://doi.org/10.1016/j.plrev.2021.07.001>.
- [6] Bizzarri M, Pontecorvi P. Critical transition across the Waddington landscape as an interpretative model: comment on “Dynamic and thermodynamic models of adaptation” by A.N. Gorban et al. *Phys Life Rev* 2021;38:115–9. <https://doi.org/10.1016/j.plrev.2021.05.010>.
- [7] Nazarenko T, Blyuss O, Whitwell H, Zaikin A. Ensemble of correlation, parenclitic and synolytic graphs as a tool to detect universal changes in complex biological systems: comment on “Dynamic and thermodynamic models of adaptation” by A.N. Gorban et al. *Phys Life Rev* 2021;38:120–3. <https://doi.org/10.1016/j.plrev.2021.05.009>.
- [8] Di Paola L, Leitner DM. Network models of biological adaptation at the molecular scale: comment on “Dynamic and thermodynamic models of adaptation” by A.N. Gorban et al. *Phys Life Rev* 2021;38:124–6. <https://doi.org/10.1016/j.plrev.2021.05.008>.
- [9] Vasenina E, Kataoka R, Loenneke JP, Buckner SL. Exercise science perspective: comment on “Dynamic and thermodynamic models of adaptation” by Alexander N. Gorban et al. *Phys Life Rev* 2021;38:129–31. <https://doi.org/10.1016/j.plrev.2021.05.005>.
- [10] Zinovyev A. Adaptation through the lens of single-cell multi-omics data: comment on “Dynamic and thermodynamic models of adaptation” by A.N. Gorban et al. *Phys Life Rev* 2021;38:132–4. <https://doi.org/10.1016/j.plrev.2021.05.004>.
- [11] Allahverdyan AE. Energy dissipation and storage in adaptation and homeostasis: comment on “Dynamic and thermodynamic models of adaptation” by Alexander N. Gorban et al. *Phys Life Rev* 2021;38:137–9. <https://doi.org/10.1016/j.plrev.2021.05.002>.
- [12] Selvarajoo K. Searching for unifying laws of general adaptation syndrome: comment on “Dynamic and thermodynamic models of adaptation” by Alexander N. Gorban et al. *Phys Life Rev* 2021;38:97–9. <https://doi.org/10.1016/j.plrev.2021.04.001>.
- [13] Giuliani A. The statistical mechanics of life: comment on “Dynamic and thermodynamic models of adaptation” by Alexander N. Gorban et al. *Phys Life Rev* 2021;38:100–2. <https://doi.org/10.1016/j.plrev.2021.04.003>.
- [14] Liu R, Aihara K, Chen L. Collective fluctuation implies imminent state transition: comment on “Dynamic and thermodynamic models of adaptation” by Alexander N. Gorban et al. *Phys Life Rev* 2021;38:103–7. <https://doi.org/10.1016/j.plrev.2021.04.002>.
- [15] Red'ko VG. Some aspects of adaptation and evolution: comment on “Dynamic and thermodynamic models of adaptation” by Alexander N. Gorban et al. *Phys Life Rev* 2021;38:108–10. <https://doi.org/10.1016/j.plrev.2021.04.004>.
- [16] Ueltzhöffer K, Da Costa L, Friston KJ. Variational free energy, individual fitness, and population dynamics under acute stress: comment on “Dynamic and thermodynamic models of adaptation” by Alexander N. Gorban et al. *Phys Life Rev* 2021;38:111–5. <https://doi.org/10.1016/j.plrev.2021.04.005>.
- [17] Sedov KR, Gorban AN, Petushkova (Smirnova) EV, Manchuk VT, Shalamova EN. Correlation adaptometry as a method of screening of the population. *Vestn Akad Med Nauk SSSR* 1988;10:69–75. PMID: 3223045.
- [18] Gorban AN, Pokidysheva LI, Smirnova EV, Tyukina TA. Law of the minimum paradoxes. *Bull Math Biol* 2011;73(9):2013–44. <https://doi.org/10.1007/s11538-010-9597-1>.
- [19] Gorban AN, Smirnova EV, Tyukina TA. Correlations, risk and crisis: from physiology to finance. *Physica A* 2010;389(16):3193–217. <https://doi.org/10.1016/j.physa.2010.03.035>.
- [20] Krasnenko AN, Pokidysheva EV, Veretnova KY, Tyukina TA. Analysis of correlations in the Russian banking system in adapting to the economic crisis of 2007–2008. *J Sib Fed Univ Math Phys* 2010;3(4):521–32.
- [21] Pokidysheva L, Ignatova I. Principal component analysis used in estimation of human's immune system, suffered from allergic rhinosinusopathy complicated with clamidiosis or without it. In: Kountchev R, Iantovics B, editors. *Advances in intelligent analysis of medical data and decision support systems. Studies in computational intelligence*, vol. 473. Heidelberg: Springer; 2013. p. 147–56.
- [22] Gorban AN, Tyukina TA, Smirnova EV, Pokidysheva LI. Evolution of adaptation mechanisms: adaptation energy, stress, and oscillating death. *J Theor Biol* 2016;405:127–39. <https://doi.org/10.1016/j.jtbi.2015.12.017>.
- [23] Karmanova IV, Razzhevaikin VN, Shpitonkov MI. Application of correlation adaptometry for estimating a response of herbaceous species to stress loadings. *Dokl Bot Sci* 1996;346–348:4–7. 1996. Translated from *Doklady Akademii Nauk SSSR*.
- [24] Giuliani A. Statistical mechanics of gene expression networks: increasing connectivity as a response to stressful condition. *Adv Syst Biol* 2014;3:1–4.
- [25] Bernardini C, Censi F, Lattanzi W, Barba M, Calcagnini G, Giuliani A, et al. Network genes in the skeletal muscle of amyotrophic lateral sclerosis patients. *PLoS ONE* 2013;8(2):e57739. <https://doi.org/10.1371/journal.pone.0057739>.
- [26] Censi F, Giuliani A, Bartolini P, Calcagnini G. A multiscale graph theoretical approach to gene regulation networks: a case study in atrial fibrillation. *IEEE Trans Biomed Eng* 2011;58(10):2943–6. <https://doi.org/10.1109/TBME.2011.2150747>.
- [27] Censi F, Calcagnini G, Bartolini P, Giuliani A. A systems biology strategy on differential gene expression data discloses some biological features of atrial fibrillation. *PLoS ONE* 2010;5(10):e13668. <https://doi.org/10.1371/journal.pone.0013668>.
- [28] Scheffer M, Carpenter SR, Lenton TM, Bascompte J, Brock W, Dakos V, et al. Anticipating critical transitions. *Science* 2012;338(6105):344–8. <https://doi.org/10.1126/science.1225244>.
- [29] Razzhevaikin VN, Shpitonkov MI. The model of correlation adaptometry and its use for estimation of obesity treatment efficiency. *Russ J Numer Anal Math Model* 2012;26(6):565–74.
- [30] Shchelkunova TA, Morozov IA, Rubtsov PM, Bobryshev YV, Sobenin IA, Orekhov AN, et al. Lipid regulators during atherogenesis: expression of LXR, PPAR, and SREBP mRNA in the human aorta. *PLoS ONE* 2013;8(5):e63374. <https://doi.org/10.1371/journal.pone.0063374>.
- [31] Siggiridou E, Kugiumtzis D, Kimiskidis VK. Correlation networks for identifying changes in brain connectivity during epileptiform discharges and transcranial magnetic stimulation. *Sensors* 2014;14(7):12585–97. <https://doi.org/10.3390/s140712585>.

- [32] Cramer AOJ, van Borkulo CD, Giltay EJ, van der Maas HLJ, Kendler KS, Scheffer M, et al. Major depression as a complex dynamic system. *PLoS ONE* 2016;11(12):e0167490. <https://doi.org/10.1371/journal.pone.0167490>.
- [33] Mojtahedi M, Skupin A, Zhou J, Castano IG, Leong-Quong RYY, Chang H, et al. Fate decision as high-dimensional critical state transition. *PLoS Biol* 2016;14(12):e2000640. <https://doi.org/10.1371/journal.pbio.2000640>.
- [34] Rybnikov SR, Rybnikova NA, Portnov BA. Public Fears in Ukrainian society: are crises predictable? *Psychol Dev Soc* 2017;29(1):98–123. <https://doi.org/10.1177/0971333616689398>.
- [35] Heiberger RH. Predicting economic growth with stock networks. *Physica A* 2018;489:102–11. <https://doi.org/10.1016/j.physa.2017.07.022>.
- [36] Longin F, Solnik B. Is the correlation in international equity returns constant: 1960–1990? *J Int Money Financ* 1995;14(1):3–26. [https://doi.org/10.1016/0261-5606\(94\)00001-H](https://doi.org/10.1016/0261-5606(94)00001-H).
- [37] Onnela J-P, Kaski K, Kertész J. Clustering and information in correlation based financial networks. *Eur Phys J B* 2004;38:353–62. <https://doi.org/10.1140/epjb/e2004-00128-7>.
- [38] Chen L, Liu R, Liu ZP, Li M, Aihara K. Detecting early-warning signals for sudden deterioration of complex diseases by dynamical network biomarkers. *Sci Rep* 2012;2:1–8. <https://doi.org/10.1038/srep00342>.
- [39] Gorban AN. Slow relaxations and bifurcations of omega-limit sets of dynamical systems. PhD thesis in Physics & Math (Differential Equations & Math Phys). Kuibyshev, Russia. 1980.
- [40] Gorban AN, Cheresiz VM. Slow relaxations of dynamical systems and bifurcations of ω -limit sets. *Dokl Akad Nauk SSSR* 1981;261(5):1050–3. Communicated by S.L. Sobolev, English translation: *Sov Math Dokl* 1981;24(3):645–9.
- [41] Gorban AN. Singularities of transition processes in dynamical systems: qualitative theory of critical delays. *Electron J Differ Equ, Monogr* 2004;05. <https://ejde.math.txstate.edu/Monographs/05/gorban.pdf>.
- [42] Gorban AN. Singularities of transient processes in dynamic and beyond. Comment on “Long transients in ecology: theory and applications” by Sergei Petrovskii et al. *Phys Life Rev* 2020;32:46–9. <https://doi.org/10.1016/j.pprev.2019.12.002>.
- [43] Kleinbub JR, Palmieri A, Orsucci FF, Andreassi S, Musmeci N, Benelli E, et al. Measuring empathy: a statistical physics grounded approach. *Physica A* 2019;526:120979. <https://doi.org/10.1016/j.physa.2019.04.215>.
- [44] Vasenina E, Kataoka R, Buckner SL. Adaptation energy: experimental evidence and applications in exercise science. *J Trainol* 2020;9:66–70. https://doi.org/10.17338/trainology.9.2_66.
- [45] Albergante L, Mirkes E, Bac J, Chen H, Martin A, Faure L, et al. Robust and scalable learning of complex intrinsic dataset geometry via EIPiGraph. *Entropy* 2020;22(3):296. <https://doi.org/10.3390/e22030296>.
- [46] EIPiGraph in R on GitHub. <https://github.com/sysbio-curie/EIPiGraph.R>, 2018.
- [47] EIPiGraph in MATLAB on GitHub. <https://github.com/sysbio-curie/EIPiGraph.M>, 2020.
- [48] EIPiGraph in Python on GitHub. <https://github.com/sysbio-curie/EIPiGraph.P>, 2020.
- [49] Chen H, Albergante L, Hsu JY, Lareau CA, Lo Bosco G, Guan J, et al. Single-cell trajectories reconstruction exploration and mapping of omics data with STREAM. *Nat Commun* 2019;10(1):1903. <https://doi.org/10.1038/s41467-019-09670-4>.
- [50] Aynaud MM, Mirabeau O, Gruel N, Grossetete S, Boeva V, Durand S, et al. Transcriptional programs define intratumoral heterogeneity of Ewing sarcoma at single-cell resolution. *Cell Rep* 2020;30(e6):1767–79. <https://doi.org/10.1016/j.celrep.2020.01.049>.
- [51] Zinovyev A, Sadovsky M, Calzone L, Fouché A, Groeneveld CS, Chervov A, et al. Modeling progression of single cell populations through the cell cycle as a sequence of switches. *bioRxiv*. <https://doi.org/10.1101/2021.06.14.448414>, 2021.
- [52] Zinovyev A, Sadovsky M, Calzone L, Fouché A, Groeneveld CS, Chervov A, et al. Python notebooks for the article “Modeling progression of single cell populations through the cell cycle as a sequence of switches”. GitHub. https://github.com/auranic/CellCycleTrajectory_SegmentModel, 2021.
- [53] Whitwell HJ, Bacalini MG, Blyuss O, Chen S, Garagnani P, Gordileva SY, et al. The human body as a super network: digital methods to analyze the propagation of aging. *Front Aging Neurosci* 2020;12:136. <https://doi.org/10.1016/j.pprev.2021.03.001>.
- [54] Sanchez R, Newman DE. Self-organized criticality and the dynamics of near-marginal turbulent transport in magnetically confined fusion plasmas. *Plasma Phys Control Fusion* 2015;57(12):123002. <https://doi.org/10.1088/0741-3335/57/12/123002>.
- [55] Landau LD. On the problem of turbulence. *Dokl Akad Nauk SSSR* 1944;44:339–42. Engl. transl. in: *C R Acad Sci URSS* 1944;44:311–4, and in Landau L.D. *Collected papers*. Oxford/New York: Pergamon Press/Gordon and Breach; 1965. p. 387–91. <https://doi.org/10.1016/b978-0-08-010586-4.50057-2>.
- [56] Hartley K, Jarvis DS. Let nine universities blossom: opportunities and constraints on the development of higher education in China. *High Educ Res Dev* 2021;1–5. <https://doi.org/10.1080/07294360.2021.1915963>.
- [57] Tilman D. Resources: a graphical-mechanistic approach to competition and predation. *Am Nat* 1980;116(3):362–93. <https://doi.org/10.1086/283633>.
- [58] Tyukin I. *Adaptation in dynamical systems*. NY: Cambridge University Press; 2011.
- [59] Bauer ES. *Theoretical biology*. Moscow: VIEM; 1935. Reprinted: Budapest: Akadémiai Kiadó; 1984.
- [60] Horowitz JM, Zhou K, England JL. Minimum energetic cost to maintain a target nonequilibrium state. *Phys Rev E* 2017;95(4):042102. <https://doi.org/10.1103/PhysRevE.95.042102>.
- [61] Gorban AN. *Equilibrium encircling. Equations of chemical kinetics and their thermodynamic analysis*. Novosibirsk: Nauka Publ.; 1984.
- [62] Gorban AN, Kaganovich BM, Filippov SP, Keiko AV, Shamansky VA, Shirkalin IA. *Thermodynamic equilibria and extrema: analysis of attainability regions and partial equilibria*. Berlin-Heidelberg-New York: Springer; 2006.