SCIENTIFIC Publishing Limited

Digital Technologies Research and Applications

http://ojs.ukscip.com/index.php/dtra

Article

Quantifying System-Environment Synergistic Information by Effective Information Decomposition

Mingzhe Yang [1], Linli Pan [1] and Jiang Zhang [1,2,*]

[1] School of Systems Science, Beijing Normal University, Beijing 100875, China

[2] Swarma Research, Beijing 102300, China

[*] Correspondence: zhangjiang@bnu.edu.cn

Received: 3 August 2025; **Revised:** 27 August 2025; **Accepted:** 3 September 2025; **Published:** 29 September 2025

Abstract: Living systems maintain structural and functional stability while adapting to environmental changes, a capability independent of specific system-environment states. Existing frameworks, such as self-organization theory and free energy principles, cannot measure system-environment interaction at the causal level. In this article, we propose a new causal indicator, Flexibility, to measure a system's ability to respond to its environment. We construct this indicator based on information theory and interventional operations from causal inference, which implies the indicator depends only on the dynamical causal mechanism. We show this indicator satisfies the axiom system of the partial information decomposition (PID) framework and decomposes into two components, Expansiveness and Introversion, which correspond to different strategic tendencies for environmental adaptation. This decomposition reveals that Flexibility depends on the entanglement between system-environment variables and noise magnitude. Through experiments on cellular automata (CA), random Boolean networks, and real gene regulatory networks (GRNs), we validate that the indicator identifies the most complex and computationally capable CA (Langton's parameter at 0.5), while demonstrating that feedback loops carrying important biological functions in GRNs exhibit the highest flexibility. We also find that flexibility peaks at a moderate level of dynamical noise. Furthermore, we combine this framework with machine learning techniques to demonstrate its applicability when the underlying dynamics are unknown.

Keywords: Synergy; Flexibility; Effective Information; Partial Information Decomposition; Gene Regulatory Networks

1. Introduction

Many complex systems exhibit life-like characteristics. Examples range from flexible robots that adapt to human interactions to self-organizing online communities that produce innovative crowdsourced outcomes. So what are the primary distinctions between them and ordinary systems? The self-organization theory first addresses this question [1]. People believe a life-like system must have the ability of self-organization [2,3]. However, life is more than just an open system with self-organization [4]. More importantly, they can maintain the stability of their own structure and function when faced with diverse and changing environments. For example, a snowflake is a self-organizing system [3], but it is not a living system. Once the environmental temperature rises, the snowflake melts. If this snowflake could autonomously avoid high-temperature environments and phase transitions that would damage its own structure, it would be an adaptive system with life activity [4]. Many papers have proposed this kind of

adaptive system in various fields [5–7].

Besides qualitative discussions, we need a formal framework to quantify and identify this unique property of life. Information theory is an important instrument for quantifying living systems [8,9]. Many metrics can measure a system's self-organization [10,11]. However, to date, few indicators exist that measure the extent to which a system can cope with environmental changes. The proposal of the information theory of individuality [12] is to measure the individuality of living systems from the perspective of information dynamics. It describes the individual survival of a system as maximizing the transmission of its own information over time, so mutual information and conditional mutual information are used to define the individuality of an organism.

In the information theory of individuality, the calculation of the metric (mutual information) depends on the state distribution of the observed data, which is contingent upon the system's initial conditions, the environment, and the dynamical process duration (if the system is non-steady-state). Conversely, the characteristic of life's flexible response to the environment is not a property that varies with time and state, but rather reflects the characteristics of the interaction mechanisms between the system and the environment, representing a dynamical property [13]. Therefore, we should focus on the quantities defined on their causal mechanisms, rather than the states [14].

Tononi and Hoel et al. have proposed information indicators measured at the causal level, such as effective information (EI) [14,15]. The causal mechanism remains invariant with respect to state and time. Here, dynamics are described by Markovian transition probability matrices (TPM) [14,16], and EI is designed as a function of TPM [17,18]. EI is used to measure the strength of causal effects in dynamics [14,18], and the difference in EI between macro and micro dynamics can quantify the emergence in complex systems [14,16]. When dynamical mechanisms are unknown, machine learning techniques can identify causal mechanisms from the data [19]. Since EI does not account for the interactions between the system and the environment [20], we extend EI to develop a causal metric that incorporates environmental influences. Consequently, we can leverage this kind of indicator, along with other relevant metrics [21], to describe causal properties and craft an indicator capable of measuring the system's responsiveness to the environment at the causal level.

Moreover, through partial information decomposition (PID) [22], we see that individuality involves redundant information and synergistic information. The information theory of individuality uses the PID theory to explain the physical meaning of individuality indicators, but fails to calculate information atoms that describe the system's flexible response to the environment [12]. Researchers have proposed numerous methods to calculate information atoms, but their computational results conflict with the consensus regarding specific properties [23]. In this paper, we introduce EI, which requires the input variables from the previous moment to be uniformly distributed [14,15], and starting from the existing PID axiomatic system [24], we derive a computable definition of information atoms in a three-variable system. We extend effective information (EI) within the PID framework to decompose system-environment interactions into three components: synergistic (flexibility), individual, and external driving information. This approach isolates the synergistic information produced by their entangled dynamics.

In this article, we define an indicator at the causal level to characterize the ability of a system to maintain its own structure and function when dealing with environmental changes. Through mathematical proofs, we show that this indicator is precisely the synergistic information of the system itself when the system and the environment are coupled together. In our numerical experiments, we demonstrate the correlation between this indicator and the type of cellular automata. Additionally, we apply this indicator to gene regulatory networks (GRNs) [25], uncovering the significance of feedback loops (FBLs) and the responses of the steady states of a highly synergistic system to environmental alterations. Notably, FBLs are instrumental in carrying out the biological functions of biological systems in reacting to environmental changes [26]. Additionally, we conducted a machine learning experiment to demonstrate that causal mechanisms can be identified using machine learning when data alone is available. Our experiments reveal that flexibility peaks for complex types of cellular automata,and at their maximal computational capability, identify feedback loops as biological flexibility hubs.

2. Formulation

Next, we provide the formal expressions for the system and the environment. Three variable combinations affect the system's state at the next time step: the system, the environment, and their combined variables. Based on this, we define three causal mechanisms: the Individual Mechanism, the External Driving Mechanism, and the Distinct Mechanism.

2.1. Distinct Mechanism

We define System X as a subset of the World U, with respective state sets Ω_X and Ω_U. Assuming the dynamics of U are discrete and Markovian, the conditional probability $P(U^{t+1}|U^t)$ characterizes them, where t represents time in the stochastic process. To exclusively measure the dynamical properties, we eliminate the influence of the World data distribution by introducing the do-operator [21], denoted as $do(U^t \sim \mathcal{U}(\Omega_U))$, where $\mathcal{U}$ represents the uniform distribution.

For the stochastic process at time t + 1, our analysis focuses exclusively on the temporal evolution of the target system X^{t+1}. This necessitates marginalizing over extraneous variables in the global system U, thereby restricting attention to the marginalized conditional probability: $P(X^{t+1}|U^t)$. Formally, this probability measure is obtained through state space projection: $\forall x^{t+1} \in \Omega_X,\ P(X^{t+1} = x^{t+1}|U^t) = \sum_{u^{t+1} \in \Omega_U,\ \pi_X(u^{t+1}) = x^{t+1}} P(U^{t+1} = u^{t+1}|U^t)$.

Here, $\pi_X : \Omega_U \to \Omega_X$ represents the canonical projection mapping that extracts the component from the global state. For example, consider a universe U with two systems, A and B, each having two states, 0 and 1. The entire universe U thus has four possible states: 00, 01, 10, and 11. To compute the transition probability for system A alone, we sum the conditional probabilities of U transitioning to 00 and 01 to obtain the probability of A transitioning to 0. Similarly, we sum the probabilities of U transitioning to 10 and 11 to obtain the probability of A transitioning to 1.

Following the terminology proposed by Albantakis et al. [27], we collectively term the system and environment as a Distinction. We are solely concerned with the impact of the system and environment on the system's state at the next moment; thus, we define the Distinct TPM as the following:

$$P_{X^t,E^t \to X^{t+1}} : P\left(X^{t+1} \middle| X^t, E^t\right) \tag{1}$$

The state spaces of the joint variable X, E, and of X are denoted as $\Omega(X, E)$ and Ω_X, respectively. Consequently, the shape of this TPM is $|\Omega(X, E)| \times |\Omega_X|$. In **Figure 1b**, we focus on the causal arrows from X^t and E^t pointing to X^{t+1}, which corresponds to the dynamics described by the system's distinct TPM.

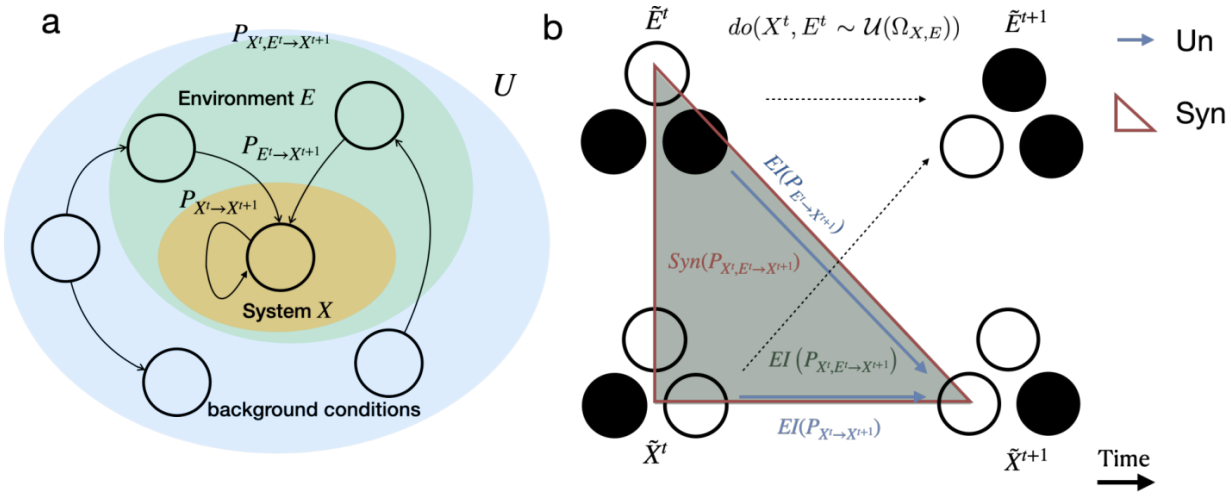

Figure 1. The causal diagram of the system's interaction with the environment. (**a**) When a set of variables is designated as the system, the spatial distinction between the system, environment, and background conditions is made, with arrows representing the causal relationships between variables. (**b**) The causal diagram of the interaction between X and E over time. Changes in the filling of circles in the diagram represent state transitions. In the diagram, the green triangle and the formula represent the EI from both the system and the environment to the system itself, the blue lines and the formulas represent the two types of unique information, and the red lines and the formula represent the effective synergistic information between the system and its environment, also known as flexibility. This diagram is consistent with the causal diagram described by Krakauer et al. [12].

2.2. Individual and External Driving Mechanisms

Starting with the distinct TPM $P_{X^t,E^t\to X^{t+1}}$, we can derive the TPM $P_{X^t\to X^{t+1}}$, which considers only the system's internal dynamics and excludes environmental information, termed the Individual Mechanism. Meanwhile, we extract the TPM $P_{E^t\to X^{t+1}}$, which focuses exclusively on the environment's impact on the system's information, termed the External Driving Mechanism. The definitions of these mechanisms are as follows, and their corresponding relationships with the causal diagrams are also marked in **Figure 1a**.

$$P_{X^t\to X^{t+1}} = |\Omega_E|^{-1} \sum_{E^t=e^t} P\left(X^{t+1}\middle|X^t, E^t\right) \tag{2}$$

$$P_{E^t\to X^{t+1}} = |\Omega_X|^{-1} \sum_{X^t=x^t} P\left(X^{t+1}\middle|X^t, E^t\right) \tag{3}$$

When we observe the probability distribution of environmental states, the conditional probability of the system can be obtained through $P\left(X^{t+1}\middle|X^t\right) = \sum_{E^t=e^t} P(E^t)P\left(X^{t+1}\middle|X^t, E^t\right)$. Because we introduce the do-operator [21], which intervenes in both the environment and the system to achieve a uniform distribution, denoted as $do(X^t, E^t \sim \mathcal{U}(\Omega_{X,E}))$ in **Figure 1b**, we have $P(E^t) = |\Omega_E|^{-1}$. Consequently, we derive the expressions for $P_{X^t\to X^{t+1}}$ and, similarly, for $P_{E^t\to X^{t+1}}$ as Equations (2) and (3) describe.

2.3. Definition of Flexibility

After obtaining those TPMs, we introduce the EI function for an arbitrary causal mechanism (TPM). EI represents the effective information, which quantifies the strength of causal effects for a TPM [14,17]. EI is defined as the mutual information for an intervened uniformly distributed input variable X and its corresponding output variable Y as shown in the following formula.

$$\begin{aligned} EI\left(P_{X\to Y}\right) &\equiv I\left(X, Y\middle|do\left(X \sim \mathcal{U}(\Omega_X)\right)\right) \\ &= \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N} p_{ij}\log\frac{N\,p_{ij}}{\sum_{k=1}^{N} p_{kj}} \end{aligned} \tag{4}$$

p_{ij} is an element of the TPM $P_{X\to Y}$, and N is the number of states of the input variable. For further details, see **Supplementary Materials Appendix B**. We can define EIs for the mechanisms $P_{X^t\to X^{t+1}}$ and $P_{E^t\to X^{t+1}}$, which are also coined as Individual Driving Information and External Driving Information:

$$EI\left(P_{X^t\to X^{t+1}}\right) = I\left(X^t, X^{t+1}\middle|do\left(X^t, E^t \sim \mathcal{U}(\Omega_{X,E})\right)\right) \tag{5}$$

$$EI\left(P_{E^t\to X^{t+1}}\right) = I\left(E^t, X^{t+1}\middle|do\left(X^t, E^t \sim \mathcal{U}(\Omega_{X,E})\right)\right) \tag{6}$$

These equations show the information about the system's next moment, provided solely by the system, and that provided solely by the environment. Naturally, due to the properties of mutual information, both types of effective information are non-negative. Within the effective joint mutual information, after the subtraction of these two components, what remains is the information that is exclusively provided by the system and the environment in union, termed as flexibility, or the effective synergy between system and environment, denoted by $Syn(P_{X^t,E^t\to X^{t+1}})$.

$$Syn\left(P_{X^t,E^t\to X^{t+1}}\right) = EI\left(P_{X^t,E^t\to X^{t+1}}\right) - EI\left(P_{X^t\to X^{t+1}}\right) - EI\left(P_{E^t\to X^{t+1}}\right) \tag{7}$$

The correspondence between these indicators and the causal diagram is presented in **Figure 1b**. We also define flexibility for continuous dynamical systems. Please refer to **Supplementary Materials Appendix F**.

2.4. Properties of Flexibility

Actually, $Syn(P_{X^t,E^t\to X^{t+1}})$ aligns with the definitions and axiomatic system of PID theory [28] regarding the requirements for synergistic information in trivariable systems.

In the following paragraphs, we denote $\tilde{Z}$ as the intervened version of any random variable Z after the intervention $do(X^t, E^t \sim \mathcal{U}(\Omega_{X,E}))$. Consequently, we have the following theorem:

Theorem 1. *In a trivariable system, the flexibility defined in Equation (7) is the synergistic information of $\tilde{X}^t$, $\tilde{E}^t$ with respect to $\tilde{X}^{t+1}$.*

This theorem pertains to a specific PID axiomatic system (please refer to **Supplementary Materials Appendix A**). For the proof of this theorem, please refer to **Supplementary Materials Appendix C**. The upper bound of synergy is $min\{I(\tilde{X}^t;\tilde{X}^{t+1}|\tilde{E}^t),\ I(\tilde{E}^t;\tilde{X}^{t+1}|\tilde{X}^t)\}$. The proof of this property is provided in **Supplementary Materials Appendix C**. We can further decompose the synergy term, i.e., Equation 7.

Corollary 1. The flexibility defined in Equation (7) can be decomposed into two components: Expansiveness and Introversion.

We define Expansiveness, abbreviated as Exp, and Introversion, abbreviated as Int, as follows:

$$Exp\left(P_{X^t,E^t\to X^{t+1}}\right) = \frac{1}{|\Omega_E|}\sum_{e\in\Omega_E} H\left(\frac{1}{|\Omega_X|}\sum_{x\in\Omega_X} P_{x,e}\right) + \frac{1}{|\Omega_X|}\sum_{x\in\Omega_X} H\left(\frac{1}{|\Omega_E|}\sum_{e\in\Omega_E} P_{x,e}\right) \tag{8}$$

$$Int\left(P_{X^t,E^t\to X^{t+1}}\right) = 2\log_2|\Omega_X| - \frac{1}{|\Omega_{X,E}|}\sum_{e\in\Omega_E}\sum_{x\in\Omega_X} H\left(P_{x,e}\right) - H\left(\frac{1}{|\Omega_{X,E}|}\sum_{e\in\Omega_E}\sum_{x\in\Omega_X} P_{x,e}\right) \tag{9}$$

Here, $P_{x,e}$ denotes the conditional probability distribution corresponding to system state x and environment state e in a distinct TPM. The proof of Corollary 1 can be found in **Supplementary Materials Appendix C**.

To elucidate the meanings of expansiveness and introversion, we first introduce the EI function and its decomposition [16].

$$EI(P_{X\to Y}) = \underbrace{-\left(\frac{1}{N}\sum_{i=1}^{N} H(P_i)\right)}_{\text{determinism}} + \underbrace{H\left(\frac{1}{N}\sum_{i=1}^{N} P_i\right)}_{\text{non-degeneracy}} \tag{10}$$

Here, X denotes an arbitrary input variable with a state space of size N, Y represents the corresponding output variable, P_i corresponds to the i-th row of the TPM $P_{X\to Y}$, and $H(P)$ indicates the Shannon entropy of the probability distribution P. The term $-\left(\frac{1}{N}\sum_{i=1}^{N} H\left(P_i\right)\right)$ represents determinism; when it is high, it indicates that the system has low noise. The term $H\left(\frac{1}{N}\sum_{i=1}^{N} P_i\right)$ represents non-degeneracy; when it is low (degeneracy is high), it implies that the system will deterministically converge to certain states, indicative of attractor dynamics.

When the system employs different TPMs corresponding to individual environmental states, the environmental context becomes explicitly incorporated. In contrast, when environmental states remain unspecified, the system adopts an environment-averaged TPM. Through the EI decomposition framework, Equation (8) measures the system's state-specific differentiation under well-defined environments versus its stochastic variability under environmental uncertainty. These dual aspects together constitute Expansiveness as the system's outward adaptive orientation. Correspondingly, Equation (9) captures the system's structured coordination in explicit environments versus its reduced differentiation in ambiguous contexts, jointly characterizing Introversion as the internal consolidation tendency. This demonstrates two distinct pathways through which the system enhances flexibility: expansiveness and introversion. The relationship between the magnitudes of Exp and Int under different conditions can be referred to in **Figure 2**. Given that the Shannon entropy of the system's probability distribution ranges from 0 to $\log_2|\Omega_X|$, we can determine the numerical ranges for Exp and Int: $0 \le Exp\left(P_{X^t,E^t\to X^{t+1}}\right),\ Int\left(P_{X^t,E^t\to X^{t+1}}\right) \le 2\log_2|\Omega_X|$.

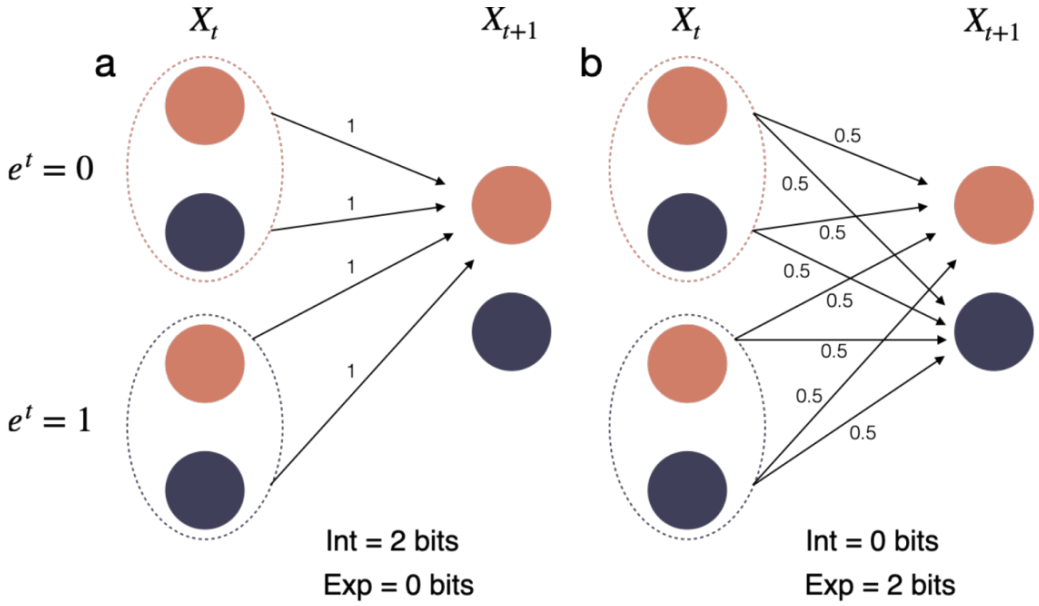

Figure 2. Schematic diagrams of expansiveness and introversion. Different colors of the circles represent different system states, while different colors of the dashed lines indicate different environmental states at that time. Arrows with associated numerical values denote transition probabilities from system-environment states at time t to system states at time t + 1. (**a**) shows the case of low expansiveness and high introversion, while (**b**) shows the case of high expansiveness and low introversion.

2.5. Summary of Terms

To clarify the meaning and physical significance of each term, we provide a summary in **Table 1**. As the table shows, Flexibility decomposes into Expansiveness and Introversion, corresponding to species with different strategic orientations. For instance, Daphnia reproduces asexually in favorable environments but switches to sexual reproduction in unfavorable ones, even if this introduces deleterious mutations (noise) [29]. It is therefore a system with high Expansiveness. In contrast, the hedgehog opts for low-risk physical protection, representing a system with high Introversion. Individual Driving Information and External Driving Information, however, fail to reflect the unique properties of living systems.

According to PID theory, Individual Driving Information and External Driving Information are unique information after an intervention: each relates to one source variable and not the other. Therefore, they cannot reflect the interaction between the two source variables (the system and the environment at the same time). Flexibility, however, corresponds to synergistic information post-intervention, which is transmitted only when considering both source variables. It can thus capture the system-environment interaction. **Supplementary Materials Appendix A** provides more details on PID theory.

Table 1. Summary of Terms.

Terms	Dynamics Description	Physical Interpretation or Examples
Flexibility	TPMs differ across environments and exhibit low noise.	Adapts to environments with diverse persistence strategies.
Expansiveness	TPMs differ across environments.	Daphnia switch reproductive modes across different environments.
Introversion	TPMs exhibit low noise.	Hedgehogs' spiny shells offer physical protection against most threats.
Individual Driving Information	The system's state depends only on its state at the previous time step.	Most inorganic substances maintain stable structures and states.
External Driving Information	The system's state depends only on the environment's state at the previous time step.	Most proteins lose activity in high-temperature environments.

3. Results

In the subsequent experiments, we will validate the meaning and functionality of these metrics.

3.1. Cellular Automaton

Cellular automata (CA), particularly one-dimensional elementary CA, are used to simulate artificial life, with Wolfram identifying 256 possible rule sets [30]. Wolfram classified CA into four behavior types: stable, periodic, chaotic, and complex, with Class IV potentially being computationally universal. However, this classification is subjective, leading Langton to introduce the parameter $\lambda = 1 - \frac{n}{8}$ [2] to quantify CA behaviors, where n is the number of outputs being 1 in the CA rule table. He demonstrated that CA behaviors can continuously transition from Class I to Class III as λ varies from 0 to 1, with values between 0.3 and 0.6 corresponding to complex Class IV dynamics. For λ around 0.5, CAs exhibit complex behavior and perform best on computational tasks [31,32].

As depicted in **Figure 3**, we have validated the relationship between flexibility and its decomposition with the behavior types of CA. In **Figure 3a**, as CA becomes increasingly complex, expansiveness rises while introversion declines, overall reflecting an increase in flexibility. A similar phenomenon is observed in **Figure 3b**. When λ is between 0.3 and 0.6, the increase in expansiveness exceeds the decrease in introversion, leading to an increase in flexibility. The following mathematical relationship indicates that, in noise-free CA, introversion is indeed a definite function of λ .

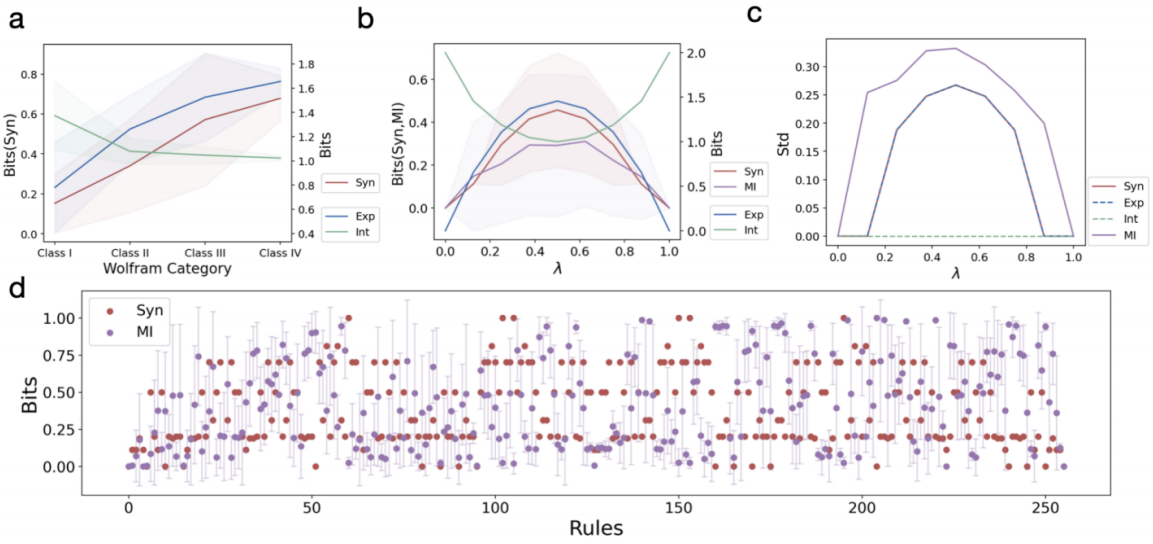

Figure 3. (**a**) The trend of expansiveness, introversion, and flexibility between the system and the environment as the type of CA changes from Class I to Class IV. The line represents the mean, and the band represents the standard deviation. (**b**) A comparative trend chart of them, along with the mutual information proposed by Langton [2], as λ varies. (**c**) A comparative chart of the standard deviation for the four indicators is calculated. (**d**) A scatter plot of flexibility and mutual information for all 256 CA. The flexibility of a specific rule-based cellular automaton is calculated without a standard deviation, whereas mutual information includes a standard deviation, depicted by error bars.

$$
\begin{aligned}
Int\left(P_{X^t,E^t\to X^{t+1}}\right) &= 2\log_2|\Omega_X| - \frac{1}{|\Omega_{X,E}|}\sum_{e\in\Omega_E}\sum_{x\in\Omega_X} H\left(P_{x,e}\right) - H\left(\frac{1}{|\Omega_{X,E}|}\sum_{e\in\Omega_E}\sum_{x\in\Omega_X} P_{x,e}\right) \\
&= 2\log_2|\Omega_X| - H\left(\frac{1}{|\Omega_{X,E}|}\sum_{e\in\Omega_E}\sum_{x\in\Omega_X} P_{x,e}\right) \\
&= 2\log_2|\Omega_X| - H(\lambda, 1-\lambda)
\end{aligned}
\tag{11}
$$

This explains why, in **Figure 3c**, the standard deviation of introversion is zero, while the fluctuations in effective synergy are attributed to expansiveness. Concurrently, it is evident that mutual information, as a measure based on observational data, exhibits a larger standard deviation, even though its trend aligns with that of effective synergy. In **Figure 3d**, we run iterations for 200 steps under each initial condition in a space of 10-cell automata. Based on these observational data, we calculate the mutual information from time t to t+1 for a single cell and compute the expectation and standard deviation over all possible initial conditions. The varying standard deviations across different rules of CA indicate that the selection of initial conditions is crucial for the computation of mutual information for some rules. In contrast, the calculation of effective synergy is independent of initial conditions.

3.2. Identify Flexible Motifs in Gene Regulatory Networks

We will measure the flexibility of gene regulatory networks (GRNs) using real data.

GRNs describe how a collection of genes governs key processes within a cell, which are often modeled as Boolean networks. Kadelka et al. [33] established the most comprehensive repository of expert-curated Boolean GRN models to date, encompassing both structural configurations and Boolean functions. These models describe the regulatory logic underlying a variety of processes in numerous species across multiple kingdoms of life. Due to computational constraints, our analysis was limited to a subset of these networks. We selected 63 models, with node counts ranging from 5 to 67, encompassing animal, plant, fungal, and bacterial domains.

To investigate which GRN structures exhibit enhanced environmental responsiveness in real-world settings, we assessed the flexibility of various three-node subgraph configurations in **Figure 4a**. The analysis revealed that feedback loops (FBLs) demonstrated the highest flexibility values. In fact, FBLs in GRNs carry important biological functions. For example, FBL structures often exhibit dynamical compensation (DC), which is the ability of a model to compensate for variation in a parameter [26]. Additionally, many oscillators in biological systems originate from negative FBLs, known as repressilators [26]. They play a crucial role in adapting to environmental changes and regulating their own cycles. In **Figure 4b**, the mean flexibility of FBLs is also high, confirming the above conclusions. Moreover, the structure with the highest value among four-node structures is not a simple FBL but a more connected structure that includes FBLs (see the highlighted part in the figure). This structure has not yet been fully studied and named by biologists. Perhaps it carries some interesting functions related to biological adaptation to the environment that have yet to be discovered. We further compare the flexibility of FBLs and FFLs with different Boolean functions; see **Supplementary Materials Appendix D** for the analysis.

To verify that systems with higher flexibility have a stronger ability to respond to environmental changes, we compared the evolution of gene activation states under environmental shocks in **Figure 4c,d**. The genes in **Figure 4c** are from the GRN of macrophage activation, with a flexibility of 0.566, while those in **Figure 4d** are from the GRN of tumour cell invasion and migration, with a flexibility of 0. Comparing the time series curves, the former exhibits a greater diversity of steady states under different environmental conditions. We measured more precisely the mean mutual information between consecutive moments when environmental state changes lead to shifts in system steady states, finding a positive correlation coefficient of 0.487 with flexibility. For further experimental details, see **Supplementary Materials Appendix D**.

3.3. Random Boolean Network and Machine Learning

To explore dynamical characteristics influencing flexibility beyond network topology, subsequent experiments investigate the effects of dynamical parameters on random Boolean networks (RBNs) with fixed structures. Furthermore, while previous experiments were conducted under known dynamical mechanisms, practical scenarios frequently involve data-driven problems with unknown underlying mechanisms. We therefore employ RBN simulations integrated with machine learning methodologies. By first reconstructing governing mechanisms from observational data and subsequently performing flexibility measurements, we validate the applicability of our framework to systems with hidden dynamical rules.

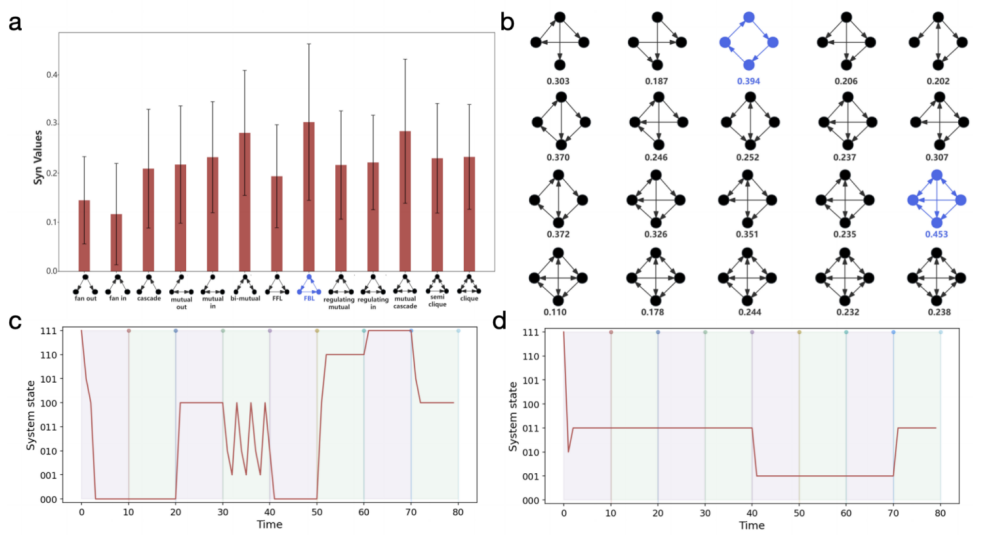

Figure 4. (**a**) The mean and standard deviation of flexibility for all possible three-node subgraph structures in various real environments, where the value for FBL is the highest. (**b**) The mean flexibility for some four-node subgraph structures, with the two highlighted structures having the highest values. (**c**) The time series of state changes for the system composed of the genes BAG4, BAG4 TNFRSF1A, and TNF BAG4 TNFRSF1A in the GRN of macrophage activation, starting from the initial state "111" and switching environmental states every 10 steps, with a total of 8 randomly selected environmental states. (**d**) The time series generated by the system composed of the genes CDH1, CDH2, and GF in the GRN of tumour cell invasion and migration, under experimental conditions consistent with those in (**c**).

In **Figure 5**, each variable can take on two values, 0 or 1. For any variable X_i in the system, its update rule is defined as follows:

$$P\left(X_i^{t+1} = 1 \middle| u^t\right) = \frac{1}{1 + \exp\left(-k \sum_{j=1}^{n} w_{j,i} u_j^t\right)} \tag{12}$$

Each variable U_j's edge acting on another variable X_i respectively, corresponds to a weight value $w_{j,i} \in [0, 1]$. $k \in [0, +\infty]$ is a parameter controlling the noise magnitude. When $k = 0$, the noise is the greatest; that is, regardless of the values of the input variables, the conditional probability is a uniform distribution. The larger k is, the smaller the noise intensity is. We set the temperature $T = \frac{1}{k}$, for $k \neq 0$. In this experiment, a total of two variables were set. One is the temperature T, and the other is the proportion of the system's own variables and the environmental variables' effect on the system. We set the weight of the system's own effect as $w \in [0, 1]$ (the solid line in **Figure 5**, including self-loops), and the weights of the effects of the environmental variables on the system (the dashed arrows in **Figure 5**) are $w_{E_1,i} = 0.5(1 - w)$ and $w_{E_2,i} = 1.5(1 - w)$. It can be seen that the larger w is, the stronger the influence of the system on itself is, and the weaker the influence of the environment on the system is. The heatmap of the flexibility varying with T and w is shown in **Figure 6a**.

As illustrated in **Figure 6a**, the system exhibits near-zero flexibility when either intrinsic self-influence or environmental influence dominates the dynamics. A maximum flexibility value emerges at optimal coupling strength w, demonstrating the critical role of balanced interactions. Furthermore, increasing the temperature T generally induces a monotonic reduction in flexibility across the parameter space. Notably, **Figure 6b** reveals a counterintuitive phenomenon where moderate noise levels ($T \approx 0.2$) paradoxically enhance flexibility to peak values. The phenomenon of enhanced synergistic information under low noise levels has been previously captured by other metrics [34]. Our experiments demonstrate that this phenomenon originates from interactions at the dynamical mechanism level. More importantly, through the decomposition of flexibility, we reveal that systems leverage low noise to achieve greater diversity and environmental sensitivity (Exp). The benefits of this trade-off outweigh the loss of intrinsic order (Int) caused by noise, resulting in an optimal noise level that maximizes flexibility. In biological systems, noise inherent in the interactions of genes regulating circadian rhythms maintains oscillatory behavior

without decay [26]. This suggests that other complex systems may benefit similarly from controlled noise levels, enabling optimal environmental responsiveness.

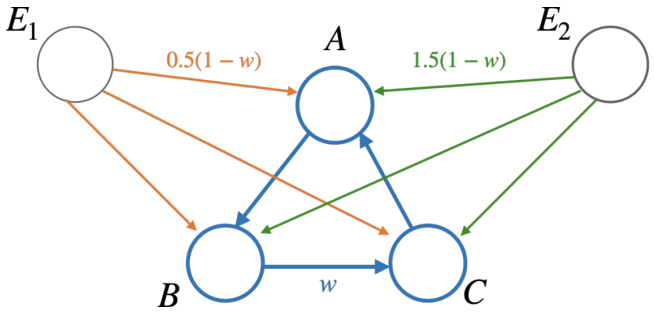

Figure 5. The schematic illustrates the experimental design framework, where nodes A, B, and C constitute the core system interacting with environmental variables E_1 and E_2. Edges indicate interaction relationships, with weights $w_{j,i}$ encoded by color-coded mathematical symbols (colored circles denote self-loops). The continuously adjustable parameter $w \in [0,\ 1]$ governs interaction intensities, whose functional role is defined through the mathematical formulation in Equation (12).

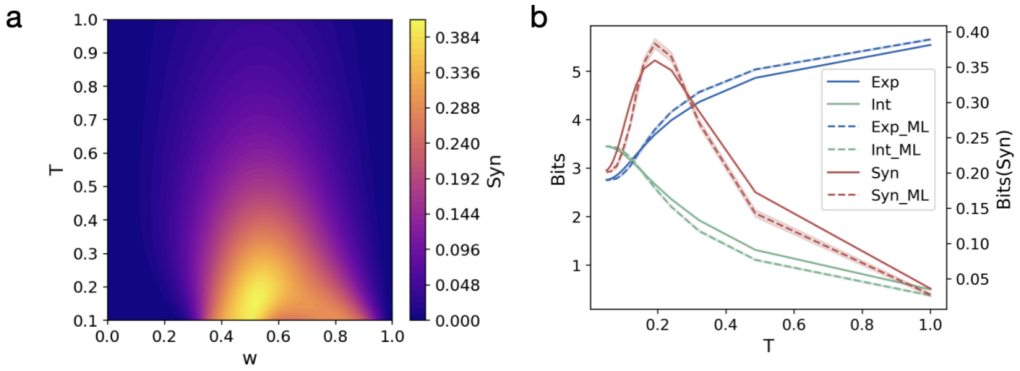

Figure 6. The experimental result graph. In (**a**), it indicates the heatmap of the flexibility varying with the temperature T and the weight ratio parameter w. (**b**) shows that when $w = 0.5$, the changing trends of flexibility, expansiveness, and introversion with respect to T. The solid line shows the ground-truth result, which corresponds to the line with $w = 0.5$ in (**a**). The dashed line is the calculation result of the trained artificial neural network based on machine learning from the generated data. The radius of the band is the standard deviation of the results of 10 repeated experiments.

To validate the framework's capability in reconstructing and quantifying system mechanisms under unknown dynamics, we train a neural network (NN) with four fully-connected layers ($32 \to 64 \to 64 \to 16 \to 8$) using data generated from conditional probabilities at $w = 0.5$ under uniform input distribution. The NN architecture employs LeakyReLU activations in hidden layers and cross-entropy loss for one-hot encoded inputs. The alignment between predicted (dashed) and theoretical (solid) curves in **Figure 6b** demonstrates successful extension of our measurement framework to data-driven scenarios through machine learning integration. We also calculate the machine learning results when the data contain varying proportions of observational noise; refer to **Supplementary Materials Appendix E** for details.

4. Discussion

This framework still needs improvement. We assumed the dynamics are Markovian, but many real-world problems require a non-Markovian framework. Future work could explicitly incorporate non-Markovian dynamics by, for example, modeling system memory or delayed feedback. Additionally, although we obtained indicators of multivariate information decomposition with good properties on the three-variable system with two variables acting on

one variable, in the future, the calculation problem of synergistic information when the source variables reach three or more needs to be addressed. In this paper, the calculation and experiments of the indicators are based on discrete systems. Having proposed a method to calculate flexibility in continuous dynamical systems (**Supplementary Materials Appendix F**), our framework can be applied to a broader range of problems within such systems.

We currently assume that the dynamics are known. If the dynamics are unknown but the data is accessible, machine learning techniques can also be used to obtain the underlying dynamic mechanism first and then measure it. The machine learning in this paper is a preliminary attempt to prove its feasibility. In the future, we can introduce more complex neural network models to learn more complex dynamic mechanisms, and even take flexibility as the optimization goal to train artificial models with higher flexibility.

Nowadays, powerful causal inference and machine learning methods can infer GRNs from multi-omics data [35]. With these data and inference methods, we can apply flexibility to identify key structures in many future practical applications, such as drug discovery [36,37] and artificial systems with computational capabilities like online learning and education [38,39].

5. Conclusions

Overall, for systems and environments that satisfy the Markov dynamics assumption, we defined on the TPM how to measure the synergistic influence of the system and the environment on the system's flexibility, which captures the flexibility of the system. It is neither completely determined by the dynamics of the system itself (different from Individual Driving Information) nor completely determined by the dynamics of the environment on the system (different from External Driving Information), but corresponds to the part in the dynamics of the system and the environment as a whole where the whole is greater than the sum of the parts. In the experiments of CA, we verified that complex cellular automata have higher flexibility. More importantly, on the Boolean network data of GRNs selected by experts, we found that the structure of feedback loops in various real environments has higher flexibility. This indicates that flexibility specifically points to the biological functions carried by this structure, such as dynamic compensation, biological cycle regulation, and so on [26]. Currently, we have only compared different Boolean network structures and have not distinguished different Boolean functions on the same structure. In the future, flexibility can be used to predict whether new structures and new dynamic functions have the biological functions we are interested in.

Through the decomposition of flexibility into dual components, we established expansiveness and introversion. These respectively measure the interaction variety between system and environment, and the level of dynamic organization in behavioral patterns. As shown in the machine learning experiments, the noise intensity in the dynamics is inversely proportional to the magnitude of introversion. Expansiveness measures, apart from the noise factor, quantify the extent to which the influences on the system from the system itself and from the environment cannot be decoupled. Combining the machine learning experiments on RBNs and the computational results on CAs, we found that when there is noise in the dynamics, the reduction of noise increases flexibility by increasing the magnitude of introversion. And when the noise in the dynamics remains unchanged, the coupling degree of the influences of the system and the environment on the system is reflected by expansiveness, and at this time, the change of flexibility is dominated by the change of expansiveness. The analysis of expansiveness indicates that in the process of obtaining the variable space from the state space through a certain partition, the known boundary between the system and the environment is only one of several possible partitions. Synergy occurs when the original boundary is too ambiguous, so that we need to find a new coarse-graining of the state space.

Supplementary Materials

The supporting information can be downloaded at https://ojs.ukscip.com/files/DTRA-1492-Supplementary-Materials.docx.

Author Contributions

Conceptualization, M.Y. and J.Z.; methodology, M.Y.; software, M.Y. and L.P.; validation, M.Y. and L.P.; formal analysis, M.Y. and J.Z.; investigation, M.Y.; resources, J.Z.; data curation, L.P. and M.Y.; writing—original draft preparation, M.Y., J.Z. and L.P.; writing—review and editing, M.Y., J.Z. and L.P.; visualization, M.Y. and L.P.; supervision, M.Y. and J.Z.;

project administration, M.Y. and J.Z. All authors have read and agreed to the published version of the manuscript.

Funding

This work received no external funding.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments

We wish to acknowledge the support of Swarma Research and the assistance of Lifei Wang, a scientist from Swarma.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Kauffman, S.A. *The Origins of Order: Self-Organization and Selection in Evolution*; Oxford University Press: Oxford, UK, 1993.
2. Langton, C.G. Computation at the edge of chaos: Phase transitions and emergent computation. *Physica D* **1990**, *42*, 12–37.
3. Lansing, J.S. Complex adaptive systems. *Annu. Rev. Anthropol.* **2003**, *32*, 183–204.
4. Friston, K.J.; Stephan, K.E. Free-energy and the brain. *Synthese* **2007**, *159*, 417–458.
5. Holland, J.H.; Miller, J.H. Artificial adaptive agents in economic theory. *Am. Econ. Rev.* **1991**, *81*, 365–370.
6. Holland, J.H. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*; MIT Press: Cambridge, MA, USA, 1992.
7. Holland, J.H. Complex adaptive systems. *Daedalus* **1992**, *121*, 17–30.
8. Tkačik, G.; Bialek, W. Information processing in living systems. *Annu. Rev. Condens. Matter Phys.* **2016**, *7*, 89–117.
9. Bartlett, S.; Eckford, A.W.; Egbert, M.; et al. The physics of life: Exploring information as a distinctive feature of living systems. *PRX Life* **2025**, *3*, 037003.
10. Rosas, F.; Mediano, P.A.; Ugarte, M.; et al. An information-theoretic approach to self-organisation: Emergence of complex interdependencies in coupled dynamical systems. *Entropy* **2018**, *20*, 793.
11. Gershenson, C.; Fernández, N. Complexity and information: Measuring emergence, self-organization, and homeostasis at multiple scales. *Complexity* **2012**, *18*, 29–44.
12. Krakauer, D.; Bertschinger, N.; Olbrich, E.; et al. The information theory of individuality. *Theory Biosci.* **2020**, *139*, 209–223.
13. Barandiaran, X.E.; Di Paolo, E.; Rohde, M. Defining agency: Individuality, normativity, asymmetry, and spatio-temporality in action. *Adapt. Behav.* **2009**, *17*, 367–386.
14. Hoel, E.P.; Albantakis, L.; Tononi, G. Quantifying causal emergence shows that macro can beat micro. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 19790–19795.
15. Oizumi, M.; Albantakis, L.; Tononi, G. From the phenomenology to the mechanisms of consciousness: Integrated information theory 3.0. *PLoS Comput. Biol.* **2014**, *10*, e1003588.
16. Hoel, E.P. When the map is better than the territory. *Entropy* **2017**, *19*, 188.

17. Yuan, B.; Zhang, J.; Lyu, A.; et al. Emergence and causality in complex systems: A survey of causal emergence and related quantitative studies. *Entropy* **2024**, *26*, 108.
18. Zhang, J.; Tao, R.; Leong, K.H.; et al. Dynamical reversibility and a new theory of causal emergence based on SVD. *npj Complex.* **2025**, *2*, 3.
19. Yang, M.; Wang, Z.; Liu, K.; et al. Finding emergence in data by maximizing effective information. *Natl. Sci. Rev.* **2024**, *12*, nwae279.
20. Varley, T.F. Flickering emergences: The question of locality in information-theoretic approaches to emergence. *Entropy* **2022**, *25*, 54.
21. Pearl, J.; Mackenzie, D. *The Book of Why: The New Science of Cause and Effect*; Basic Books: New York, NY, USA, 2018.
22. Williams, P.L.; Beer, R.D. Nonnegative decomposition of multivariate information. *arXiv* **2010**, arXiv:1004.2515.
23. Lizier, J.T.; Bertschinger, N.; Jost, J.; et al. Information decomposition of target effects from multi-source interactions: Perspectives on previous, current, and future work. *Entropy* **2018**, *20*, 307.
24. Bertschinger, N.; Rauh, J.; Olbrich, E.; et al. Shared information—new insights and problems in decomposing information in complex systems. In *Proceedings of the European Conference on Complex Systems 2012*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 251–269.
25. Ghosh, D.; Marwan, N.; Small, M.; et al. Recent achievements in nonlinear dynamics, synchronization, and networks. *Chaos* **2024**, *34*, 063132.
26. Alon, U. *An Introduction to Systems Biology: Design Principles of Biological Circuits*; Chapman and Hall/CRC: Boca Raton, FL, USA, 2019.
27. Albantakis, L.; Barbosa, L.; Findlay, G.; et al. Integrated information theory (IIT) 4.0: Formulating the properties of phenomenal existence in physical terms. *PLoS Comput. Biol.* **2023**, *19*, e1011465.
28. Williams, P.L. *Information Dynamics: Its Theory and Application to Embodied Cognitive Systems*; Ph.D. Thesis, Indiana University, Bloomington, IN, USA, 2011.
29. Zhang, Y.-N.; Zhu, X.-Y.; Wang, W.-P.; et al. Reproductive switching analysis of *Daphnia similoides* between sexual female and parthenogenetic female by transcriptome comparison. *Sci. Rep.* **2016**, *6*, 34241.
30. Wolfram, S. Statistical mechanics of cellular automata. *Rev. Mod. Phys.* **1983**, *55*, 601–644.
31. Mitchell, M.; Hraber, P.; Crutchfield, J.P. Revisiting the edge of chaos: Evolving cellular automata to perform computations. *arXiv* **1993**, adap-org/9303003. [CrossRef]
32. Teuscher, C. Revisiting the edge of chaos: Again? *Biosystems* **2022**, *218*, 104693.
33. Kadelka, C.; Butrie, T.-M.; Hilton, E.; et al. A meta-analysis of Boolean network models reveals design principles of gene regulatory networks. *Sci. Adv.* **2024**, *10*, eadj0822.
34. Orio, P.; Mediano, P.A.; Rosas, F.E. Dynamical noise can enhance high-order statistical structure in complex systems. *Chaos* **2023**, *33*, 043128.
35. Leong, R.; He, X.; Beijen, B.S.; et al. Unlocking gene regulatory networks for crop resilience and sustainable agriculture. *Nat. Biotechnol.* **2025**, *43*, 1254–1265.
36. Jiang, W.; Ye, W.; Tan, X.; et al. Network-based multi-omics integrative analysis methods in drug discovery: A systematic review. *BioData Min.* **2025**, *18*, 27.
37. Farhadi, A.; Zamanifar, A.; Faezipour, M. Application of generative AI in drug discovery. In *Application of Generative AI in Healthcare Systems*; Springer: Cham, Switzerland, 2025; pp. 155–174.
38. Dong, S.; Zhou, W. Improved influential nodes identification in complex networks. *J. Intell. Fuzzy Syst.* **2021**, *41*, 6263–6271.
39. Dong, S.; Tao, X.; Zhong, R.; et al. Advanced mathematics exercise recommendation based on automatic knowledge extraction and multilayer knowledge graph. *IEEE Trans. Learn. Technol.* **2024**, *17*, 776–793.