

Article

# Navigability, Walkability, and Perspicacity Associated with Canonical Ensembles of Walks in Finite Connected Undirected Graphs—Toward Information Graph Theory

Dimitri Volchenkov 

Department of Mathematics and Statistics, Texas Tech University, 1108 Memorial Circle, Lubbock, TX 79409, USA; dimitri.volchenkov@ttu.edu

**Abstract:** Canonical ensembles of walks in a finite connected graph assign the properly normalized probability distributions to all nodes, subgraphs, and nodal subsets of the graph at all time and connectivity scales of the diffusion process. The probabilistic description of graphs allows for introducing the quantitative measures of navigability through the graph, walkability of individual paths, and mutual perspicacity of the different modes of the (diffusion) processes. The application of information theory methods to problems about graphs, in contrast to geometric, combinatoric, algorithmic, and algebraic approaches, can be called information graph theory. As it involves evaluating communication efficiency between individual systems' units at different time and connectivity scales, information graph theory is in demand for a wide range of applications, such as designing network-on-chip architecture and engineering urban morphology within the concept of the smart city.

**Keywords:** statistical ensembles of walks; information graph theory; navigability, walkability, and perspicacity in networking systems

**MSC:** 05Cxx; 05C82; 90Bxx; 91D30



**Citation:** Volchenkov, D. Navigability, Walkability, and Perspicacity Associated with Canonical Ensembles of Walks in Finite Connected Undirected Graphs—Toward Information Graph Theory. *Information* **2023**, *14*, 338. <https://doi.org/10.3390/info14060338>

Academic Editors: Anirban Bandyopadhyay and Kanad Ray

Received: 4 May 2023  
Revised: 8 June 2023  
Accepted: 12 June 2023  
Published: 15 June 2023



**Copyright:** © 2023 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction to Information Graph Theory

Real-world networking systems comprising functional units that interact with each other at different and disparate temporal and spatial scales are often metaphorically represented by finite connected undirected graphs [1]. In graph models, the vertices embody the system functional units, and the edges sketch the various interactions between these units. Graph models are useful for computer networking, sociology, urban morphology, smart city networks, chemistry, engineering, communication, logistics and management, security applications, etc., as they help us to answer the fundamental question about relations between the *local* and *global* properties of a networking system.

In our work, we assess the quality of these relations by using some *information* quantities at different graph connectivity scales, ranging from the connectivity between the nearest neighbors to the connectivity with respect to infinitely long paths available in the graph. The proposed information approach to the description of graph structures (which can be called *information graph theory*) may be especially useful for the assessment of *communication efficiency* between individual systems' units at different time and connectivity scales.

One possible application of graph information theory is the design of network-on-chip (NoC), or network-based communications subsystems on an integrated circuit, most typically between semiconductor intellectual property cores schematizing various functions of the computer system in a system on a single chip [2]. (We profoundly thank our reviewer for pointing us at the possible NoC-related application of graph information theory.) NoCs provide the advantage of customized architecture, increased scalability, and bandwidth and come in many network topologies, many of which are still experimental [3].

The interconnections between multiple cores on a chip present a considerable influence on the performance and communication of the chip design regarding the throughput, end-to-end delay, and packet loss ratio. Based on the huge amount of supported heterogeneous cores, efficient communication between the associated processors has to be considered at all levels of system design to ensure global interconnection [4]. Although hierarchical architectures have addressed the majority of the associated challenges of the traditional interconnection techniques, the main limiting factor is scalability, and NoC has proven to be a promising solution. As communicating nodes require a routing algorithm for successful transmission of packets, it is important to design an NoC routing algorithm capable of providing less congested paths, better energy efficiency, and high scalability [5]. It is important to mention that the same choices of routing algorithms, when considered for more than one performance parameter of the network, do not yield expected results, since different applications exhibit different network performances for the same routing technique [6]. Fault-tolerant mapping and routing techniques at different levels of an NoC were surveyed in [7]. In [8], a method for choosing hierarchical coordinates and a greedy forwarding algorithm for path-finding were presented. Nevertheless, reliability is becoming a major concern in NoC design [7]. In particular, the accurate predictions of network packet latency, contention delay, and the static and dynamic energy consumption of the network are in demand [9].

Another evident application field of graph information theory is the study of urban morphology, i.e., urban forms and structures affecting sustainability and urban growth. Along with the spreading of new technologies and the availability of big data, cities are increasingly viewed not simply as places in space but as systems of networks and flows [10]. This became particularly evident in the concept of a *smart city*, as the synergy of architecture, technology, and the *Internet of Things* which enables the collection and data exchange of objects embedded with electronics, software, sensors, and network connectivity [11].

All above-mentioned research problems involve a variety of information quantities resolving uncertainty about evolution of some diffusion processes at different connectivity scales in a graph, and therefore the entire approach to graphs summarized in the present paper may be called *information graph theory*.

Thinking about transportation in networks—imagine the weekend pedestrian flow in a city downtown area—we may notice that some streets are more preferred by walkers than others. While some urban quarters are visited quite often, others seem to be excluded from the rest of urban fabric, as though being barricaded by invisible gates, literally turning the deserted city block into a ghetto. A widespread belief in the influence of the spatial organization of the built environment on the way people move through urban spaces and meet others by chance was common in architectural and urban thinking for centuries [12–14]. The dynamics of poverty areas in London studied over the past century in [15] has shown that the creation of poverty areas is ultimately a spatial process relating to spatial segregation and poverty. In order to quantify isolation by identifying *the most and least probable paths through the areas with respect to a certain exploration strategy*, in the present paper we put forward a concept of *walkability*.

Navigation focuses on locating a walker's position with respect to known locations, traversed paths, and structural patterns in the environment [16]. In the probabilistic setting, *navigability* would represent a *share* of predictable information about the future location of a traveler given their walking history and the present location. The predictable information about the future location of a walker is obtained with the use of two major navigation strategies, known as *path integration* (which allows for keeping track of the position and heading while exploring a new space) and *landmark-based piloting* (re-calculating position of a traveler while in a familiar environment), working in concert during navigation in humans and animals [16].

Finally, as such a circumstance appears quite often in the context of science and technology, we may pose the following question: *How long should one observe a networking system for the information collected to be enough to draw a reliable conclusion about its evolution in*

*the long run?* Of course, an answer to the above-posed question ultimately depends on the structural complexity of a system (that can often be represented by a graph), the dynamics of the processes in question, as well as on the chosen graph exploration strategy—all are difficult to quantify, in general. In the present paper, we call such a deep, penetrating discernment of the system’s nature and evolution *perspicacity* and discuss and quantify it in the framework of a proposed model.

In our work, we put forward a toy model, in which the questions of interest may receive quantitative answers. The role of a complex system will be played by a finite connected undirected graph providing a scene for *isotropic* and a variety of *anisotropic* diffusions, i.e., different types of *random walks*, forming the *canonical statistical ensembles of walks* defined in graphs (see Section 3). Uncertainty about the current state in Markov chains contains *predictable* (apparent uncertainty) and *unpredictable* (true uncertainty) information components: the conditional mutual information quantities among the past, present, and future states of the chain (see Section 2). The amount of predictable information can be viewed as a divergence between the *ensemble average* (defined with respect to the stationary distribution of random walks) and the *time-average* of walks describing the probability of observing a walker in a unit time of random walks. Consequently, the amount of uncertainty about the future location of random walkers in the graph that can be resolved by using the standard navigation strategies, such as *path integration* (based on the entire history of travels) and *landmark-based piloting* (based on the present location of the walker), can be viewed as a *navigability potential* of a node in the given graph structure with respect to the chosen exploration strategy (the type of diffusion process) (see Section 4). We call the opposite potential characterizing the amount of true uncertainty about the walker’s location in the graph *strive*, which is quantified as the expected exploration time of a node’s neighborhood by the given type of random walk (Section 4). Information flows associated with the diffusion processes in graphs (see Sections 3 and 5) can be viewed as the stochastic point processes, called *determinantal process* or fermionic processes [17], because the corresponding densities are the squared wedge-products (determinants) of eigenvectors of the symmetrized transition matrices. We use the properties of the determinantal processes to assess *perspicacity* and *walkability* in graphs (see Sections 3 and 5). The perfect perspicacity is possible only for the processes that enjoy the same type of symmetry. In Section 6, we discuss an example of the walkability analysis in urban neighborhoods and, eventually, conclude on a discussion of graph information theory and its applications in the last section.

## 2. Predictable and Unpredictable Information in Markov Chains

The resolution of uncertainty about future events is known as information [18]. For example, each coin flip reveals a single bit of information. The side repeating probability  $0 \leq p \leq 1$  in the coin tossing sequence does not affect the amount of information released at every flip, but rather determines a *share* of predictable and unpredictable information components in every released bit of information [19–22]. Indeed, the forthcoming state of the Markov chain describing tossing of a biased coin is completely predictable for  $p = 1$  when coin tossing generates the stationary sequences *heads, heads, heads,...* or *tails, tails, tails, ...*, as well as for  $p = 0$  when an alternating sequence of states arises. However, the side of the coin that will come up on the next toss is utterly unpredictable if  $p = 1/2$ .

For a Markov chain  $\{X_t\}, t \geq 0$ , defined over a finite set of  $N$  states by a (stationary) irreducible, row-stochastic transition probability matrix  $\Pr(X_{t+1} = j | X_t = i) = T_{ij} \geq 0$ , the expected amount of uncertainty about a state over infinitely long sequences is given by the Boltzmann–Gibbs–Shannon entropy [20,22–24],

$$H(X_t) = - \sum_{k=1}^N \pi_k \log_N \pi_k, \tag{1}$$

where  $\pi_k = \Pr(X_t = k)$  is a stationary distribution of states over the chain, i.e., the left eigenvector of the Markov chain transition matrix,  $\pi T = \pi$ , belonging to the major eigenvalue

1. In all information-related calculations, we shall assume that  $0 \cdot \log 0 = \log 0^0 = \log 1 = 0$ . As the information shared between the past and future in Markov chains is transmitted only through the present, the time average of the conditional entropy of the present state conditioned on all past states simplifies to the entropy rate [19,22,25,26], viz.,

$$\lim_{\tau \rightarrow \infty} \frac{1}{\tau} \sum_{t=1}^{\tau} H(X_t | X_{t-1}, \dots, X_1) = H(X_{t+1} | X_t) = - \sum_{k=1}^N \pi_k \sum_{s=1}^N T_{ks} \log_N T_{ks}. \tag{2}$$

Then, the following formal decomposition involving one-step conditional entropies [19–22],

$$\begin{aligned}
 H(X_t) &= \underbrace{\left( H(X_t) - H(X_{t+1} | X_t) \right)}_{\substack{I(X_{t+1} | X_t) \\ \text{Predictable component} \\ H(X_t | X_{t+1}, X_{t-1})}} + \underbrace{\left( H(X_{t+1} | X_{t-1}) - H(X_t | X_{t-1}) \right)}_{I(X_{t+1}, X_t | X_{t-1})} \\
 &\quad + \underbrace{\left( H(X_{t+1} | X_t) + H(X_t | X_{t-1}) - H(X_{t+1} | X_{t-1}) \right)}_{\text{Unpredictable component}}, \tag{3}
 \end{aligned}$$

represents the state’s uncertainty as a sum of the following three information components: (i)  $I(X_{t+1} | X_t)$ , expressing the amount of information about the future state that can be inferred from the entire past history of the chain, excepting for its present state; (ii)  $I(X_{t+1}, X_t | X_{t-1})$ , the mutual information between the future and present states of the chain, expressing the amount of information about the future state that can be inferred from the present state, independently of the past; and (iii)  $H(X_t | X_{t+1}, X_{t-1})$ , expressing the *true* uncertainty about the present state of the chain that cannot be resolved by observing the past history and that does not have repercussions for the future evolution of the chain. Assuming that all past states and the present state of the chain are known, we conclude that the first two quantities contribute to the predictable information component of the chain, viz.,

$$\mathcal{P}(X_t) \equiv I(X_{t+1} | X_t) + I(X_{t+1}, X_t | X_{t-1}), \tag{4}$$

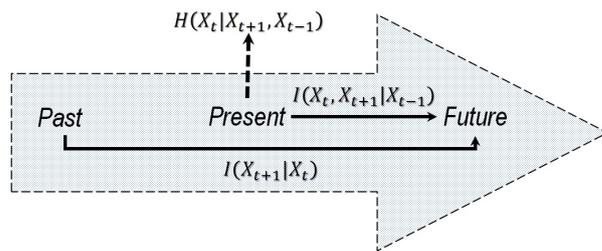
while the last term quantifies the ephemeral, unpredictable information component, viz.,

$$\mathcal{U}(X_t) \equiv H(X_t | X_{t+1}, X_{t-1}) = H(X_t) - \mathcal{P}(X_t), \tag{5}$$

which exists at the present moment only and then disappears without consequences [19,20].

The information diagram for a Markov chain is shown in Figure 1. In the context of Markov chains, we may interpret or explain a future event either as a consequence of history, or as an immediate outcome of the present event. However, for  $0 < p < 1$ , some amount of uncertainty remains unresolved before the event occurs, and such a surprise cannot be avoided, eliminated, or controlled. Given the Markov chain transition matrix  $T$ , the entropy excess  $I(X_{t+1} | X_t)$  [19–22,27–29] reads as follows:

$$I(X_{t+1} | X_t) = - \sum_{k=1}^N \pi_k \left( \log_N \pi_k - \sum_{s=1}^N T_{ks} \log_N T_{ks} \right). \tag{6}$$



**Figure 1.** information diagram for a Markov chain. The amount of uncertainty about the future state of the chain can be resolved partially by observing its past states (at an amount of  $I(X_{t+1}|X_t)$ ), as well as from the present state alone (quantified by  $I(X_{t+1}, X_t|X_{t-1})$ ). True uncertainty about the state ( $H(X_t|X_{t+1}, X_{t-1})$ ) exists at the present moment only and then disappears without consequences. All three components sum to the entropy  $H(X_t)$ , quantifying uncertainty of a state over the infinitely long sequences generated by the Markov chain.

The conditional mutual information  $I(X_{t+1}, X_t|X_{t-1})$  between the present and the future states of the chain conditioned on the past state takes the following form [19–22]:

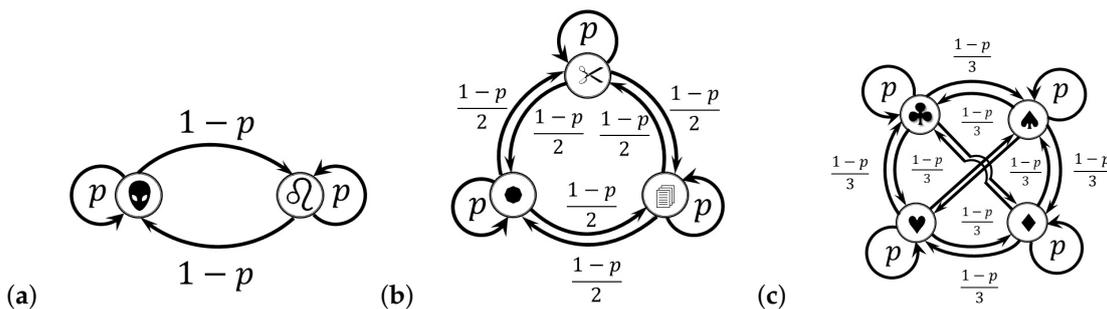
$$I(X_{t+1}, X_t|X_{t-1}) = \sum_{k=1}^N \pi_k \sum_{r=1}^N \left( T_{kr} \log_N T_{kr} - T_{kr}^2 \log_N T_{kr}^2 \right). \tag{7}$$

The conditional mutual information (7) quantifies a consequence of the present state for the future evolution of the Markov chain. Finally, the amount of true uncertainty existing only at the present moment of time is the entropy residue [19–22], viz.,

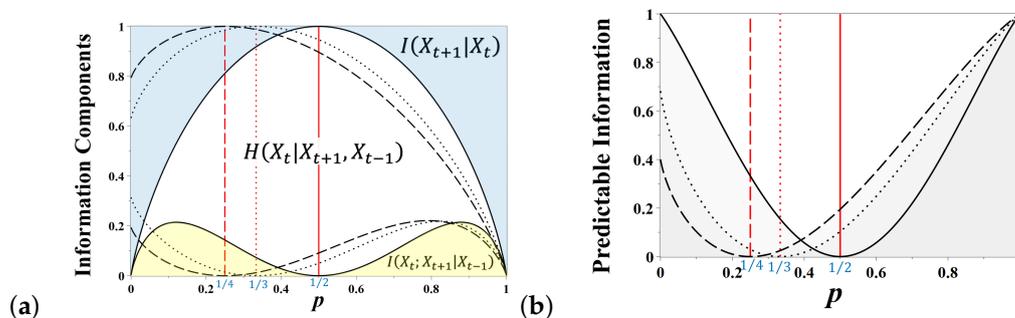
$$\begin{aligned} H(X_t|X_{t+1}, X_{t-1}) &= H(X_t) - I(X_{t+1}|X_t) - I(X_{t+1}, X_t|X_{t-1}) \\ &= - \sum_{k=1}^N \pi_k \cdot \log_2 \left[ \prod_{s=1}^N (T_{ks}^2)^{-T_{ks}^2} \right]. \end{aligned} \tag{8}$$

For example, let us consider three games of chance shown in Figure 2a–c having  $N = 2, 3$ , and 4 states, respectively. The probability of repeating a state in each game is taken as  $0 \leq p \leq 1$ . Other transitions are chosen to be equiprobable. For entropy calculations in this section, we use the  $N$ -based logarithm  $\log_N$ , so that the information and entropy quantities for these games of chance are measured in *bits*, *trits*, and *quadrants*, respectively. Namely, a single *bit* of information is released at every coin flip (see Figure 2a); one *trit* of information is produced at every round of the rock–paper–scissors game (see Figure 2b); and one *quadrant* of information is brought about at every round of the card suit game of four states (see Figure 2c). The use of different information units allows for comparing the structure of information components for all three games in a single frame (as shown in Figure 3). Otherwise, the graphs would scale up by  $\log_2 3 \approx 1.585$  and  $\log_2 4 = 2$  if we used bits as a major information unit. The detailed information analysis of a two-state Markov chain shown in Figure 2a has been developed by us in [20], and in [30] we discussed information components in a three-state chain (Figure 2b). The four-state chain shown in Figure 2c has never before been discussed.

The information components (6–8) for all three games shown in Figure 2 are presented in Figure 3 as functions of the state repetition probability  $p$ . In Figure 3a, the predictable information components in a coin tossing game (see Figure 2a) are indicated by the solid line, and the graph regions representing the predictable information components  $I(X_{t+1}|X_t)$  and  $I(X_{t+1}, X_t|X_{t-1})$  are highlighted in blue and yellow, respectively.



**Figure 2.** Three games of chance, with  $N = 2, 3$ , and  $4$  states, respectively. In each game, the probability of repeating a state is  $0 \leq p \leq 1$ . Other transitions are chosen to be equiprobable. The information components (6–8) for all three games are the functions of a single independent variable  $p$ , as shown in Figure 3. (a) The two-state game of chance describing tossing of a biased coin. (b) The three-state game of chance corresponding to the rock–paper–scissors game. (c) The four-state game of chance corresponding to the card game.



**Figure 3.** (a) The information components (6–8) as functions of the state repetition probability  $p$ . The bold line represents the information quantities for the coin tossing (see Figure 2a). The amount of information that can be predicted by observing the past history of the chain is highlighted in blue. The amount of information that can be guessed from the present state alone is highlighted in yellow. The information components for the rock–paper–scissors game (Figure 2b) are shown as dotted lines. The information components for the card suit game (Figure 2c) are shown as dashed lines. (b) The amounts of predictable information in three games of chance as functions of the state repetition probability  $p$ .

The future outcome of coin tossing is reliably predictable for  $p = 0$  or  $p = 1$  when the game generates the stationary alternating or constant sequences of states, respectively. In this case, the entire uncertainty of a state equals the entropy excess (blue in Figure 3a), viz.,  $1 \text{ bit} \equiv H(X_t) = I(X_{t+1}|X_t)$ . As  $p > 0$  or  $p < 1$ , the reliability of predictions based on the past history of the chain decays; however, the future outcome may then be guessed from the present state of the coin alone, either by alternating the present side of the coin at the next tossing (for  $p > 0$ ), or by repeating it (for  $p < 1$ ). The corresponding amount of information,  $I(X_{t+1}, X_t|X_{t-1})$ , quantifying the degree of reliability of such a guess, is highlighted by yellow in Figure 3a. When  $p \approx 1/2$ , there is a destructive interference between two incompatible hypotheses about the future outcome of coin tossing that causes the gradual attenuation and cancellation of mutual information,  $I(X_{t+1}, X_t|X_{t-1})$ , and the coin becomes fair at  $p = 1/2$ . At this moment, the entire uncertainty of a state equals the amount of true uncertainty quantified by  $H(X_t|X_{t+1}, X_{t-1})$  [20]. Summing the predictable information components together, we obtain the amount of predictable information about the future outcome of coin tossing as a function of  $p$ , shown in Figure 3b as a solid line. The coin tossing game (see Figure 2a) is predictable at  $p = 0$  and  $p = 1$ , but unpredictable at  $p = 1/2$  when the coin is fair.

The information components for other games of chance (Figure 2b,c) are presented in Figure 3a,b by the dotted and dashed lines. The rock–paper–scissors game (Figure 2b)

becomes fair at  $p = 1/3$ , and the relevant information components are shown as dotted lines. The model of a three-sided coin (*susceptible–infected–immune*) was recently used for assessing pandemic uncertainty on conditions of vaccination and self-isolation [30]. The card suit game (Figure 2c) is fair when  $p = 1/4$ , as shown by the dashed lines in Figure 3a,b. In contrast to the simple coin flipping game, other games of chance, having more possible states, are not completely predictable for  $p = 0$ , as they retain a certain degree of uncertainty of symbols in the alternating sequences of chain states.

### 3. Symmetries and Statistical Ensembles of Walks in Finite Connected Undirected Graphs

A finite connected undirected graph  $G(V, E)$ , in which  $V, |V| = N$ , is a set of vertices and  $E \subseteq V \times V$  is a set of edges, is defined by a symmetric adjacency matrix,  $A = A^T$ , such that  $A_{ij} = 1$ , iff  $(i, j) \in E$ , but  $A_{ij} = 0$  otherwise. The adjacency matrix allows for the following spectral decomposition:  $A_{ij} = \sum_{s=1}^N \alpha_s u_{is} u_{sj}$ , in which  $\{\mathbf{u}_i\}_{i=1}^N$  is an orthogonal system of eigenvectors belonging to the ordered real eigenvalues  $\alpha_{\max} \equiv \alpha_1 > \alpha_2 \geq \dots \geq \alpha_N$ , where  $\alpha_{\max}$  is simple according to the Perron–Frobenius theorem [31].

*Graph symmetries:* Multiple eigenvalues of the adjacency matrix indicate the presence of non-identity automorphisms in the graph  $G$  [19,32–36]. A group of a graph’s automorphisms  $\text{Aut}(G)$  preserving the graph structure is formed by all permutations  $\Pi \in \mathcal{P}(V)$  defined on the set of graph vertices  $V$  that commute with the graph adjacency matrix,  $[A, \Pi] \equiv A\Pi - \Pi A = 0$  [19,32,34]. If a non-identical  $\Pi \in \text{Aut}(G)$  exists, and  $\mathbf{u}$  is an eigenvector of the adjacency matrix, such that  $A\mathbf{u} = \alpha_u \mathbf{u}$ , then the vector  $\mathbf{v} = \Pi\mathbf{u}$  is another eigenvector of  $A$  belonging to the same eigenvalue  $\alpha_u$ , as  $\alpha_u \mathbf{v} = \Pi\alpha_u \mathbf{u} = \Pi A\mathbf{u} = A\Pi\mathbf{u} = A\mathbf{v}$ . An automorphism associated with some eigenvalue  $\alpha$  of multiplicity  $m_\alpha \leq (N - 1)$  is represented by a low-dimensional  $m_\alpha \times m_\alpha$  real orthogonal submatrix corresponding to the related block of a permutation matrix. As the eigenvectors of a symmetric adjacency matrix are real and mutually orthogonal,  $(\mathbf{u}, \mathbf{v}) = (\mathbf{u}, \Pi\mathbf{u}) = 0$ , it also follows that elements of these eigenvectors are  $\{0, \pm 1\}$  mapped to each other under the action of  $\Pi \in \text{Aut}(G)$ . For example, the major (Perron) eigenvector  $\langle 1, 1, 1 \rangle$  of the adjacency matrix of a triangle graph belongs to the maximum eigenvalue  $\alpha_{\max} = 2$ , while another, double eigenvalue  $\alpha_{2,3} = -1$  corresponds to the eigenvectors  $\langle -1, 1, 0 \rangle$  and  $\langle -1, 0, 1 \rangle$  mapped to each other by a simple 2-permutation.

*Microcanonical ensemble of walks:* The microcanonical ensemble of walks (MCE) is a collection of  $n$ -walks in  $G$  taken with the same probability  $P_n \equiv \exp(F_n/kT)$  for all walks of the same length  $n$ , in which the (Boltzmann constant and) temperature  $kT \equiv 1/\ln 2$ , and the free energy of  $n$ -walks is defined by the logarithm of the number of  $n$ -walks available in the graph  $G$ ,  $F_n \equiv \log_2 \sum_{ij} A_{ij}^n$ . For very long walks,  $n \gg 1$ ,  $A_{ij}^n \approx \gamma_1^2 \alpha_{\max}^n$ , where  $\gamma_1 \equiv \sum_{i=1}^N u_{i1}$ , and therefore the intensive free energy (per edge absorbed by a very long MCE walk) equals the graph topological entropy [21,22,37], viz.,

$$\mu \equiv \lim_{n \rightarrow \infty} \frac{F_n}{n} = \log_2 \alpha_{\max}. \tag{9}$$

*Canonical ensemble of walks:* In the canonical ensembles of walks (CNE) defined on  $G$ , each of  $\kappa_i^{(n)} \equiv \sum_{j=1}^N (A^n)_{ij}$  walks of length  $n$  available from every node  $i \in V$  are taken with equal probability. As  $\kappa_i^{(n+1)} = \sum_{j=1}^N A_{ij} \kappa_j^{(n)}$ , the CNE walks are defined by the following irreducible row-stochastic transition matrices [19,21,22]:

$$W_{ij}^{(n)} = \frac{A_{ij} \kappa_j^{(n)}}{\kappa_i^{(n+1)}}, \quad \kappa_i^{(0)} = 1, \quad \sum_{j=1}^N W_{ij}^{(n)} = 1, \quad n \in \mathbb{N}. \tag{10}$$

The CNE walks  $W_{ij}^{(n)}$  are the random walks defining the following stationary distributions (centrality measures [21,22]) of nodes:

$$\pi_i^{(n)} = \frac{\kappa_i^{(n)} \kappa_i^{(n-1)}}{\sum_{s=1}^N \kappa_s^{(n)} \kappa_s^{(n-1)}}, \quad \sum_{s=1}^N \pi_r^{(n)} W_{rs}^{(n)} = \pi_s^{(n)}. \tag{11}$$

In particular, the locally isotropic first order random walks, in which a random walker chooses a node to visit with equal probability among all available nearest neighbors,  $W_{ij}^{(1)} = A_{ij}/\kappa_i^{(1)}$ , defines the well-known stationary distribution of nodes (degree centrality), viz.,  $\pi_i^{(1)} = \kappa_i^{(1)}/2|E|$  [19,21,22]. For  $n \rightarrow \infty$ ,  $\kappa_i^{(n)} \approx \alpha_{\max}^n \gamma_1 u_{i1}$ , so that the series of transition matrices converges to the Ruelle–Bowen random walk [38], viz.,

$$W_{ij}^{(\infty)} = \lim_{n \rightarrow \infty} \frac{A_{ij} \kappa_j^{(n)}}{\kappa_i^{(n+1)}} = \frac{A_{ij} u_{j1}}{\alpha_{\max} u_{i1}}, \tag{12}$$

with a stationary distribution  $\pi_i^{(\infty)} = u_{i1}^2$  [39] that is related to the eigenvector centrality of nodes [40].

*Grand canonical ensemble of walks:* Finally, the grand canonical ensemble of walks (GCE) describes the statistics of “fluctuations” of the local growth rate of the number of walks available at the node  $\log_2 \kappa_i^{(n)}$  over a “chemical potential”  $n\mu$  of the equiprobable  $n$ -walks in the graph  $G$ , viz.,

$$\mathcal{P}_i^{(n)} \propto \exp\left(\frac{n\mu - \log_2 \kappa_i^{(n)}}{kT}\right). \tag{13}$$

In the thermodynamic limit of infinite walks  $n \rightarrow \infty$ , the exponential factor representing a node’s “fugacity” takes the following form [22]:

$$\lim_{n \rightarrow \infty} \exp\left(\frac{n\mu - \log_2 \kappa_i^{(n)}}{kT}\right) = \frac{1}{u_{i1} \gamma_1}, \quad \gamma_1 \equiv \sum_{j=1}^N u_{j1}. \tag{14}$$

The limiting distribution of relative fugacity, describing the change of the system of infinitely long walks (free energy of walks) due to adding/removing a node, takes the form of a Fermi–Dirac distribution, viz.,

$$\mathcal{P}_i = \lim_{n \rightarrow \infty} \mathcal{P}_i^{(n)} = \frac{1}{1 + u_{i1} \sum_{j \neq i}^N u_{j1}^{-1}}, \tag{15}$$

typical for the distribution of fermions in a single-particle state  $i$  of the energy  $\varepsilon_i \equiv \log_2(u_{i1} \sum_{j \neq i}^N u_{j1}^{-1})$ . The Fermi–Dirac distribution, previously observed only in a quantum system of non-interacting fermions, appears in the model due to the non-interacting quality of walks in the graph, similar to the Pauli exclusion principle allowing for only two possible microstates for each single-particle level [22].

*Determinantal point processes associated with canonical ensembles of walks—Perspicacity:* The CNE walks in a connected undirected graph define the determinantal point processes characterized by the probabilities of random currents calculated as the squared determinants of submatrices made up from the elements of eigenvectors of the symmetrized random walk transition matrices. These determinantal point processes provide the properly normalized probability distributions for all nodes, subgraphs, and nodal subsets at all time scales of the corresponding diffusion process [22]. Slater determinants in quantum mechanics [41] were historically the first examples of determinantal point processes that later acquired multiple applications in physics, applied mathematics, and statistics [42–45].

The symmetrization similarity transformation [19,22] makes the Markov matrix  $W_{ij}^{(n)}$  into a symmetric form,  $\widehat{W}_{ij}^{(n)}$ , viz.,

$$\widehat{W}_{ij}^{(n)} = \sqrt{\frac{\kappa_i^{(n)}}{\kappa_i^{(n+1)}}} A_{ij} \sqrt{\frac{\kappa_j^{(n)}}{\kappa_j^{(n+1)}}}, \quad \kappa_i^{(n)} \equiv \sum_{j=1}^N (A^n)_{ij}. \tag{16}$$

In the rest of this section, we omit the index  $(n)$  in the symmetrized transition matrix (16) to simplify the notations. The determinantal point processes may be defined for any type of random walk, and the resulting probability distributions indeed vary for different  $n$ . The symmetrized transition matrix (16) can be written in a spectral form as follows:  $\widehat{W} = \Phi \Lambda \Phi^\top$ , in which  $\Phi \in O(N)$  is an orthogonal matrix,  $\det \Phi \equiv \text{Sgn}(\Phi) = \pm 1$ , with columns being the orthonormal eigenvectors  $\phi_k : V \rightarrow S_1^{N-1}$  of the symmetrized transition matrix (16), and  $\Lambda$  is a diagonal matrix of real valued, ordered eigenvalues of the transition matrix,  $1 = \lambda_1 > \lambda_2 \geq \dots \geq \lambda_N \geq -1$ . The repeating eigenvalues indicate the presence of symmetries in the diffusion process, and if  $\phi_k$  is an eigenvector of the symmetrized transition matrix (16) belonging to the multiple eigenvalue  $\lambda_k$ , then there exists a permutation matrix  $\Pi$ , such that  $[\widehat{W}, \Pi] = 0$ , and  $\Pi \phi_k$  is another eigenvector of  $\widehat{W}$  belonging to the same eigenvalue  $\lambda_k$ . In the latter case, the symmetry may be represented by a low-dimensional real orthogonal submatrix in  $\Phi$  corresponding to the related blocks of a permutation matrix.

As the dynamics of discrete time random walks evolves with powers of the corresponding transition matrix, the eigenvalues  $0 < |\lambda_k| < 1$ , for  $k > 1$ , describe the rates of relaxation processes toward a stationary distribution over graph nodes belonging to the maximal eigenvalue  $\lambda_1 = 1$ . Herewith, the positive eigenvalues  $\lambda_k > 0, k > 1$ , describe the exponential decay of the relaxation processes (as  $|\lambda_k| < 1$ ), while the negative eigenvalues correspond to damped oscillations, and both fade away within the finite characteristic relaxation times  $t_k = -1 / \ln |\lambda_k|$ . The characteristic relaxation time of the stationary distribution is taken to be  $t_1 = \infty$ .

The  $k$ -th order wedge products,  $\phi_{i_1} \wedge \dots \wedge \phi_{i_k}$ , are the determinants of the corresponding  $k$ -th order minors of the matrix  $\Phi$ . The properties of determinants of submatrices composed of orthonormal vectors were discussed in [19,46] in more detail. In particular, the squared determinants of the  $k$ -th order minors of the orthogonal matrix  $\Phi$  define the properly normalized probability distributions over the index set  $\{i_1, \dots, i_k\}$ , viz.,  $|\phi_{i_1} \wedge \dots \wedge \phi_{i_k}|^2 = 1$ . For example, the squared elements of the major eigenvector of the symmetrized transition matrix,  $\phi_1^2$ , can be considered the squared determinants of the primitive submatrices of size 1, determining the stationary distribution of random walks, viz.,  $\phi_{1,i}^2 \equiv \text{Pr}_\infty(i) = \pi_i, \sum_{i \in V} \pi_i = 1$ . The corresponding probability  $\text{Pr}_\infty(i) > 0$  is interpreted as a measure of likeliness to observe a random walker at  $i \in V$  in infinite observation time  $t_1 = \infty$ . Other eigenvectors,  $\phi_k, k > 1$ , also enjoy the proper normalization,  $\phi_{k,i}^2 \equiv \text{Pr}_{t_k}(i), \sum_{i \in V} \text{Pr}_{t_k}(i) = 1$ , and therefore can be interpreted as the ephemeral probability distributions decaying in the finite characteristic time,  $t_k < \infty$ .

The squared wedge products,  $|\phi_k \wedge \phi_l|^2 = 1$ , define the ephemeral probability distributions over pairs of nodes (not necessarily corresponding to graph's edges) fostered by the modes  $\phi_k$  and  $\phi_l$  of the diffusion process, viz.,

$$\text{Pr}_{t_k, t_l}(i, j) \equiv \left| \begin{matrix} \phi_{k,i} & \phi_{k,j} \\ \phi_{l,i} & \phi_{l,j} \end{matrix} \right|^2, \quad \sum_{i,j \in V} \text{Pr}_{t_k, t_l}(i, j) = 1. \tag{17}$$

$\text{Pr}_{t_k, t_l}(i, j), i \neq j$ , may be interpreted as the joint probability mass functions (correlation functions) describing the probability amplitudes for a random walker to sojourn at the nodes  $i$  and  $j$  during the characteristic times  $t_l$  and  $t_k$ . Using the properties of determinants, we may expand the probability mass function  $\text{Pr}_{t_k, t_l}(i, j)$  into the joint probability of in-

dependent observations of the walker at  $i$  and  $j$  during the characteristic times  $t_l$  and  $t_k$ , conditional on obviously incompatible events, such as the simultaneous presence of the random walker at two different nodes, viz.,

$$\begin{aligned} \Pr_{t_k, t_l}(i, j) &= \phi_{k,i}^2 \phi_{l,j}^2 + \phi_{k,j}^2 \phi_{l,i}^2 - 2\phi_{k,i} \phi_{k,j} \phi_{l,i} \phi_{l,j} \\ &= \frac{\Pr_{t_k}(i) \Pr_{t_l}(j) + \Pr_{t_k}(j) \Pr_{t_l}(i)}{-2\sqrt{\Pr_{t_k}(i) \Pr_{t_l}(j) \Pr_{t_k}(j) \Pr_{t_l}(i)}}. \end{aligned} \tag{18}$$

If  $k = l$  and  $\lambda_k$  and  $\lambda_l$  are the simple eigenvalues, then the determinants in (17) equal zero for any  $i$  and  $j$ . Otherwise, if  $\phi_k$  and  $\phi_l$  belong to the same multiple eigenvalue, then  $\Pr_{t_k, t_l}(i, j) > 0$  for  $i$  and  $j$  involved in permutations expressing the symmetry of the diffusion process. In the latter case, the vectors  $\phi_k$  and  $\phi_l$  are localized at the graph nodes that enjoy the symmetry, while other vector elements equal zero. As the absolute value of a determinant is nothing but the volume of the parallelogram spanned by two vectors,  $\langle \phi_{k,i}, \phi_{k,j} \rangle$  and  $\langle \phi_{l,i}, \phi_{l,j} \rangle$ , for the determinant to be maximal, these vectors should be orthogonal to each other. The vectors  $\phi_k$  and  $\phi_l$  are mutually orthogonal in  $\mathbb{R}^N$ ; however, many of their elements are zero (as corresponding to the block structure of the symmetry-related permutation matrix), and others are related to each other by permutation, so there may be  $2 \times 2$  orthogonal submatrices, for which the determinant (17) is maximal.

To evaluate the amount of mutual information that knowing either of the diffusion process mode described by the eigenvector  $\phi_l$  provides about the other one described by the eigenvector  $\phi_k$ , we introduce the *perspicacity matrix* as the mutual information between the diffusion modes as follows:

$$I(\phi_k, \phi_l) = \sum_{i,j \in V} \Pr_{t_k, t_l}(i, j) \log_2 \left( \frac{\Pr_{t_k, t_l}(i, j)}{\Pr_{\infty}(i) \Pr_{\infty}(j)} \right). \tag{19}$$

As all nodes' densities are  $\Pr_{\infty}(i) > 0$  in a connected graph, the mutual information (19) is a non-negative symmetric matrix having positive elements for every pair of different eigenmodes  $k \neq l$ , with the zero elements on the major diagonal,  $k = l$ , due to the properties of the determinants. Herewith, the maximal perspicacity (mutual information) in (19)  $I(\phi_k, \phi_l)$  is obtained for the mutually orthogonal eigenvectors  $\phi_k$  and  $\phi_l$  belonging to the same multiple eigenvalues (if there are any) of Markov transition matrices (10), indicating the presence of symmetries in the diffusion processes. These symmetries depend on the chosen CNE walks (10), grounded in the structural symmetry of the graph (see Section 5 for an example).

Similarly, we may define the third- and higher-order correlation functions using the squared determinants of corresponding submatrices of  $\Phi$ .

#### 4. Navigability and Strive with Different Exploration Strategies

Random walks of every type ( $n \geq 1$ ) discussed in Section 3 introduce densities (11) over graph nodes playing the role of centrality measures [20] and defining quantitative measures of navigability for graph nodes. Frequently visited sites are predicted more efficiently than those that are little frequented, especially in the long run: the more frequent a node, the more predictable the navigator's position visiting it [21]. The two major navigation strategies working in concert during wayfinding in humans [16,47] and animals [48,49]—*landmark-based piloting* and *walk (path) integration*—might be associated with the predictable information components (7) and (6) of the Markov chains, defined by the irreducible row-stochastic transition matrices  $W^{(n)}$ ,  $n \geq 1$ , and by Equations (10) and (12), defining the CNE random walks on the graph  $G$  [21]. Namely, the (information) efficacy of the path integration strategy taking the entire walk history into account to infer the forthcoming navigator's position cannot exceed the entropy excess  $I(X_{t+1}|X_t)$  (6), and the efficacy of landmark-based piloting that uses the present location of the navigator to predict her forthcoming position can be assessed through the conditional mutual information between the present and future positions of the navigator,  $I(X_{t+1}, X_t|X_{t-1})$  (7).

As discussed in [21], the predictable (4) and unpredictable information components of Equations (5) and (8),  $\mathcal{P}(X_t) \equiv -\sum_{k=1}^N \pi_k \log_2 \varphi_k$  and  $\mathcal{U}(X_t) \equiv -\sum_{k=1}^N \pi_k \log_2 \psi_k$ , have the same form as the entropy function,  $-\sum_{k=1}^N \pi_k \log_2 \pi_k = H(X_t) = \mathcal{P}(X_t) + \mathcal{U}(X_t)$ , where we have introduced

$$\varphi_k \equiv \pi_k \cdot \prod_{s=1}^N (W_{ks}^2)^{W_{ks}^2}, \text{ and } \psi_k \equiv \prod_{s=1}^N (W_{ks}^2)^{-W_{ks}^2}, \tag{20}$$

the *navigability* and *strive* potentials of a node, such that  $\pi_k = \varphi_k \psi_k$ . The product factor in (20) has a form of *time-average* (the *geometric mean*) probability of the 2-step transition from  $k$  to  $s$  calculated over an infinitely long time of observation, viz.,

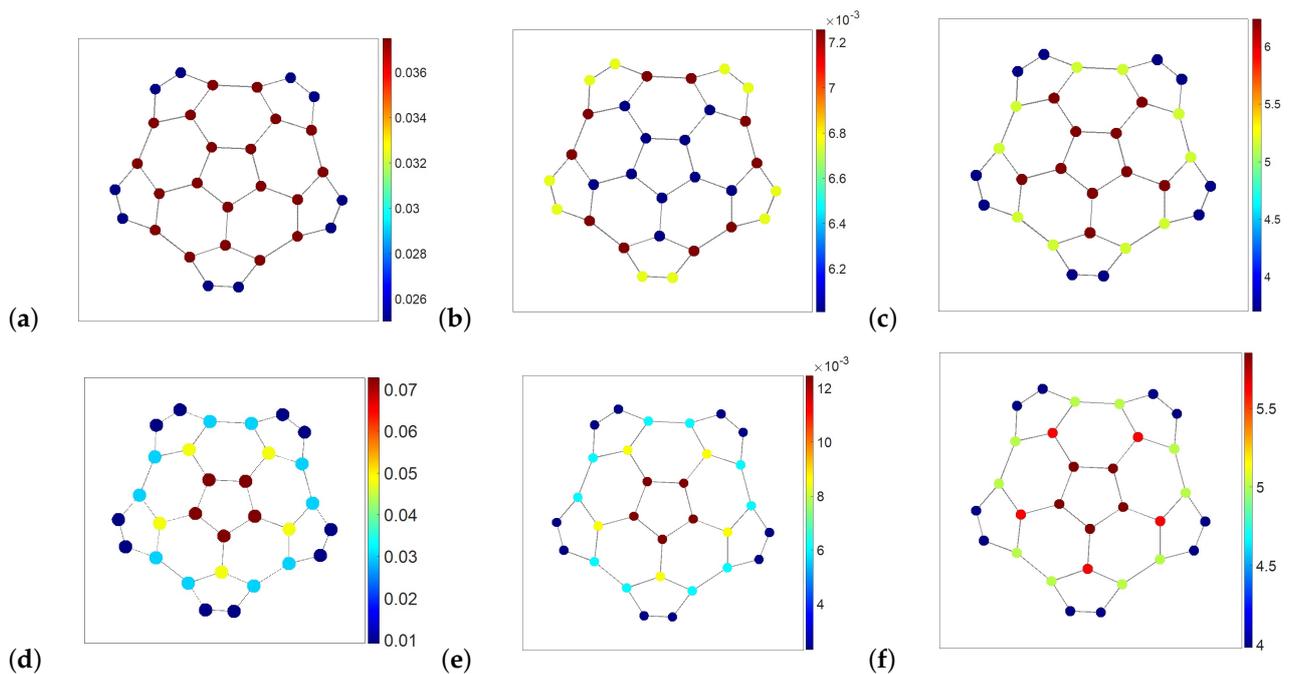
$$\lim_{\tau \rightarrow \infty} \prod_{s=1}^N \sqrt[\tau]{(W_{ks}^2)^{m_{k,s}(\tau)}} = \lim_{\tau \rightarrow \infty} \prod_{s=1}^N (W_{ks}^2)^{\frac{m_{k,s}(\tau)}{\tau}} = \prod_{s=1}^N (W_{ks}^2)^{W_{ks}^2}, \tag{21}$$

in which the number of 2-step transitions observed between the nodes  $k$  and  $s$  per time interval  $\tau$ ,  $m_{k,s}(\tau)/\tau$ , tends to the  $ks$ -element of the squared transition matrix,  $W_{ks}^2$ , as  $\tau \rightarrow \infty$ .

Given that the probability of finding a random walker at  $k$  equals  $\pi_k$ , the navigability potential  $\varphi_k$  defined accordingly (20,21) can be interpreted as the probability that the node  $k$  is a 2-step *precursor* for other nodes in the graph (including  $k$  itself) at any moment of time. Consequently, the *strive* potential  $\psi_k$  of the node  $k$  (as the inverse probability) may be interpreted as the *expected time* for the walker located at  $k$  to explore the entire 2-step neighborhood of the  $k$ -neighborhood's *exploration time*: the more sizable the neighborhood, the longer its exploration time and the stronger the strive potential, reducing the amount of predictable information about the location of the random walker.

The navigability and strive potentials of a node depend on the node's function in the graph structure, as well as on the chosen graph exploration strategy described by the random walks transition matrices  $W_{ij}^{(n)}$  (10) for some  $n \geq 1$ . In Figure 4, we presented a buckle graph with nodes color-coded according to the node's density, navigability, and strive potentials calculated for the different graph exploration strategies represented by the isotropic ( $W_{ij}^{(1)}$ ) and maximally anisotropic ( $W_{ij}^{(\infty)}$ ) random walks, respectively. Namely, in Figure 4a–c, the graph nodes are color-coded with respect to the quantities calculated for the isotropic, first-order walks, in which a random walker chooses the next node equiprobably over all nearest neighbors. The corresponding node density defined with respect to  $W_{ij}^{(1)}$  (see Figure 4a) represents the node degree centrality: the better-connected nodes have a higher probability of hosting a random walker.

The navigability potential of nodes for the random walks  $W_{ij}^{(1)}$  is shown in Figure 4b. The nodes that have fewer neighbors, as well as those connected to them (located on graph's boundary), are easy to navigate by isotropic random walks. On the contrary, the location of a walker over the better-connected nodes located at the graph center is less predictable. Conversely, the strive potential of nodes with respect to the random walks  $W_{ij}^{(1)}$  presented in Figure 4c is maximal for the central nodes of the graph (the longest neighborhood exploration time).



**Figure 4.** The color-coded density, navigability and strive potentials of nodes in a buckle graph under different graph exploration strategies. (a) The density of nodes with respect to the isotropic random walks defined by  $W_{ij}^{(1)}$  (node's degree centrality). (b) The navigability potential of nodes for the random walks  $W_{ij}^{(1)}$ . The low-connectivity nodes and nodes connected to them located on the graph's boundary are easy to navigate while the walker's position at the center of the graph is less predictable. (c) The strive potential of nodes for the random walks  $W_{ij}^{(1)}$  is maximal for the central nodes of the graph (of the longest neighborhood exploration time). (d) The density of nodes with respect to the Ruelle–Bowen random walks  $W_{ij}^{(\infty)}$  (12) of maximal anisotropy (the square root of the node's eigenvector centrality). Repelling from the graph boundaries, these random walks are anchored at the central nodes hosting the maximal share of the infinitely long paths available in the graph. (e) The navigability potential of nodes with respect to the Ruelle–Bowen random walks  $W_{ij}^{(\infty)}$  is dominated by the node's density (the square root of the node's eigenvector centrality). (f) The strive potential of nodes for the Ruelle–Bowen random walks  $W_{ij}^{(\infty)}$  (neighborhood's exploration time) decays from the graph's center to its periphery.

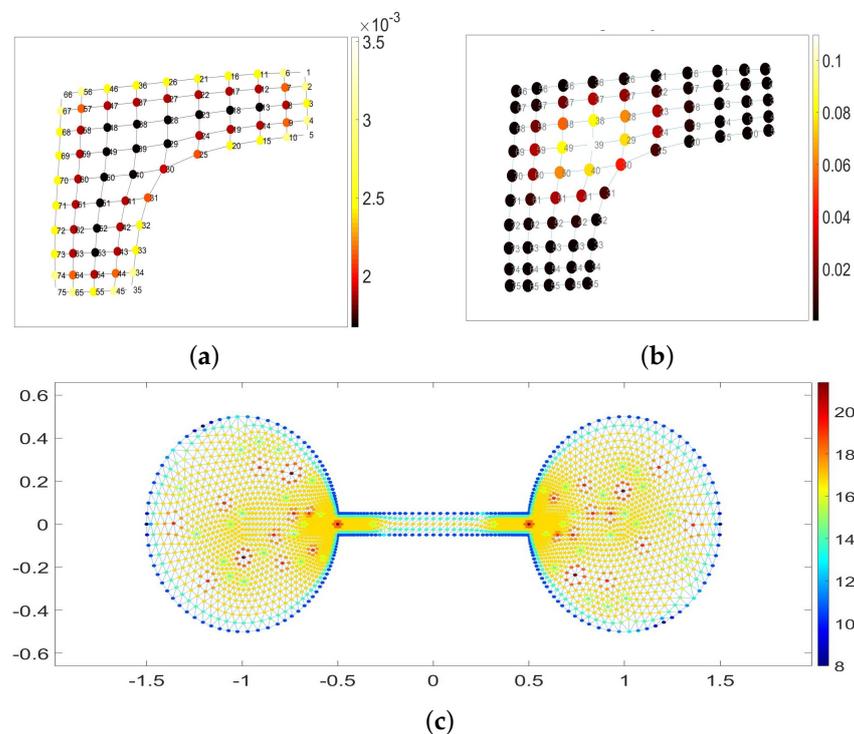
The node density with respect to the Ruelle–Bowen random walks of maximal anisotropy defined by the transition matrix  $W_{ij}^{(\infty)}$  (12) corresponds to the square root of the node's eigenvector centrality (see Figure 4d). The Ruelle–Bowen random walks are statistically confined to the graph center, hosting the lion's share of available infinite paths but repelling from the graph boundaries. The confinement of anisotropic walks at the best structurally integrated part of the graph reveals itself by the fact that the magnitude of the node's density at the graph center is as twice as high with respect to  $W_{ij}^{(\infty)}$  than with respect to  $W_{ij}^{(1)}$  and, vice versa, the Ruelle–Bowen random walkers are less probably found on the graph's boundaries than those following isotropic random walks.

The navigability potential of nodes with respect to Ruelle–Bowen random walks  $W_{ij}^{(\infty)}$  presented in Figure 4e is dominated by the node's density for anisotropic walks (the square root of the node's eigenvector centrality) and enhanced by the effect of their confinement at the graph center. The appearance of anisotropic walks in the central part of the graph is more predictable than that on the graph boundary.

In contrast to isotropic random walks, the strive potential of nodes with respect to the Ruelle–Bowen random walks  $W_{ij}^{(\infty)}$  is highest over the ring of five central nodes of the graph (see Figure 4f). The neighborhood exploration time of the center by the Ruelle–

Bowen random walks slightly exceeds five (steps), indicating that these walks are indeed statistically confined within the central ring.

In Figure 5a,b, we contrasted the navigability potentials of nodes in a membrane graph explored by isotropic random walks  $W_{ij}^{(1)}$  (a) and by Ruelle–Bowen random walks  $W_{ij}^{(\infty)}$  (b), respectively. Again, the isotropic walks are most predictable along the graph’s boundaries and at the membrane’s corners especially, but the location of walkers can hardly be predicted while in the central part of the graph. On the contrary, the navigability of the central nodes is highest if the graph is explored by Ruelle–Bowen random walks  $W_{ij}^{(\infty)}$  due to the statistical confinement of anisotropic walks at the best structurally integrated part of the graph. Finally, in Figure 5c, we have presented the strive potential of nodes (neighborhood’s exploration times) in a barbell graph explored by isotropic random walks  $W_{ij}^{(1)}$ , and the nodes of the highest degree indeed generate the maximal strive for those walks.



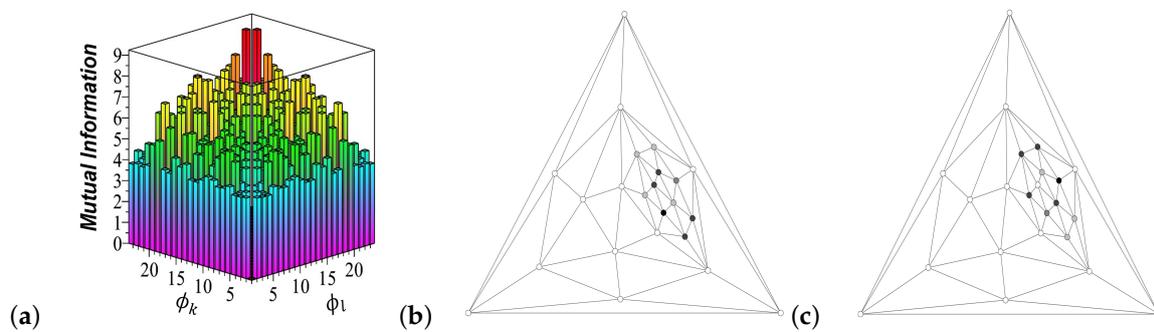
**Figure 5.** (a) The navigability potential of nodes in a membrane graph explored by Ruelle–Bowen random walks  $W_{ij}^{(\infty)}$ . (b) The navigability potential of nodes in a membrane graph explored by isotropic random walks  $W_{ij}^{(1)}$ . (c) The strive potential of nodes (neighborhood’s exploration times) in a barbell graph.

As the expected node’s neighborhood exploration time serves as a statistical estimate of its neighborhood size,  $\psi_k = \mathcal{N}_k$ , we may conclude that the navigation properties of the graph vertices are determined by the sizes of their neighborhoods with respect to the chosen graph exploration strategy. In particular, the navigability potential of a node as defined in (20) is nothing else but the “specific” density of the vertex over the size of its neighborhood, viz.,  $\varphi_k = \pi_k / \mathcal{N}_k$ . This observation can be formulated as the following “uncertainty principle” for isotropic random walks in graphs: *the larger the neighborhood of a vertex, the less reliable the assumption that a random walker is at it*. In Figure 5c, the statistical neighborhoods of the best-connected nodes are highlighted in dark red and red as being characterized by the maximal strive magnitudes with respect to the isotropic random walks in the barbell graph. The anisotropic random walks are statistically confined within the best-integrated nodes of the graph, furthest from the graph’s boundaries and structural irregularities, and therefore these nodes are the most predictable of all.

### 5. Perspicacity and Walkability of the Determinantal Point Processes Associated with Canonical Ensembles of Walks

Perspicacity, as defined in Section 3 as the mutual information (19) between the different modes of diffusion process (described by the mutually orthogonal eigenvectors of a symmetrized transition matrix of random walks), quantifies the ability to discern and understand either of the diffusion modes by knowing the other. Although perspicacity depends on the type of random walk implemented for the graph exploration, it attains its maximum value over the diffusion modes (eigenvectors) localized at the same symmetric subgraphs in the graph (if any exist) and belonging to the same repeating eigenvalue of the related Markov transition matrix.

In Figure 6a, we present the perspicacity matrix (19) calculated for the maximally anisotropic Ruelle–Bowen random walks (12) defined on the Kittel graph shown in Figure 6b,c. The Kittel graph contains a symmetric subgraph (of 9 vertices) affecting the diffusion processes that may be defined on that. For instance, the symmetrized transition matrix of Ruelle–Bowen random walks has two eigenvectors,  $\phi_{18}$  and  $\phi_{19}$ , related to each other by some permutation and belonging to the same double eigenvalue,  $\lambda = -0.358225932$ .



**Figure 6.** (a) The perspicacity matrix (19) calculated for the maximally anisotropic Ruelle–Bowen random walks (12) defined on the Kittel graph with 23 vertices and 63 edges (shown in panels (b,c)). The symmetric perspicacity matrix represents mutual information (measured in bits) between the diffusion modes defined by the eigenvectors  $\phi_k$  and  $\phi_l$ . The mutual information (perspicacity) attains its maximum value for the eigenvectors ( $\phi_{18}$  and  $\phi_{19}$ ) belonging to the same double eigenvalue ( $\lambda = -0.358225932$ ), indicating the structural symmetry in the graph. (b) The vertices of the Kittel graph are highlighted according to the squared elements  $\phi_{18}^2$  by hues of gray (white stands for 0, black stands for 1). The eigenvector  $\phi_{18}$  is localized at the graph nodes that enjoy the graph symmetry (automorphism) and that are related to the elements of the eigenvector  $\phi_{19}$  shown in (c) by permutation. (c) The vertices of the Kittel graph are highlighted according to the squared elements  $\phi_{19}^2$  by hues of gray (white stands for 0, black stands for 1). The eigenvector  $\phi_{19}$  is localized at the graph nodes that enjoy the graph symmetry (automorphism) and that are related to the elements of the eigenvector  $\phi_{18}$  shown in (b) by permutation.

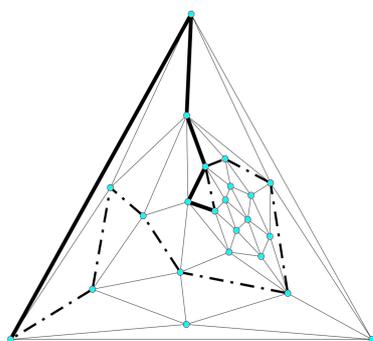
The squared elements of the eigenvectors  $\phi_{18}$  and  $\phi_{19}$  are shown in Figure 6b,c. The vertices of the Kittel graph are highlighted by hues of gray (white stands for 0, black stands for 1) according to the squared elements of  $\phi_{18}^2$  in Figure 6b and the squared elements of  $\phi_{19}^2$  in Figure 6c. The mutual information (19) between the eigenvectors  $\phi_{18}$  and  $\phi_{19}$  attains its maximum available value (of 9 bits). Thus, perspicacity, a penetrating discernment of either of the diffusion modes when the other mode is known, is fully possible only within the framework of structural symmetry common to both modes.

In urban studies, *walkability* is an essential concept of sustainable urban design associated with the ideal of a dense, mixed-used, walkable neighborhood fostering higher levels of creativity, reduced crime, improved health, and civic engagement, in contrast to the dull and deadly cul-de-sacs of car-oriented suburbs [50]. Walkability relies on the complex interdependence between built environment density, mix, and pedestrian access in synergy, which is difficult to quantify and measure for urban neighborhoods. In the present

section, we propose a mathematically robust approach to the walkability assessment in built environments (naturally represented by connected undirected graphs). Essentially, our approach is based on the following idea: *the most walkable path in a graph is that most likely to be walked* (over a relatively long period of observations). The determinantal point process associated with CNE walks can be used to estimate the likelihood that either graph edge is traversed by random walkers. Namely, each and every edge of the graph  $(i, j) \in E$  is assigned the probability (17) to be chosen for a walk during the longest available characteristic times  $t_1 = \infty$  and  $t_2$ , viz.,

$$\Pr_{t_1, t_2}(i, j) \equiv \left| \begin{array}{cc} \phi_{1,i} & \phi_{1,j} \\ \phi_{2,i} & \phi_{2,j} \end{array} \right|^2. \quad (22)$$

The standard Dijkstra's algorithm [51,52] for finding the shortest/longest paths between nodes in a weighted graph can then be used to chart the most/least probable paths between two given nodes. In Figure 7, we have highlighted the most probable (shown by the solid line) and least probable (shown by the dot-dashed line) paths between the given pair of vertices in the Kittel graph.



**Figure 7.** The most probable path (shown by the solid line) and the least probable path (shown by the dot-dashed line) between the given pair of vertices in the Kittel graph. All graph nodes are shown in blue for color contrasting.

## 6. Discussion: On the Walkability Analysis of Urban Neighborhoods

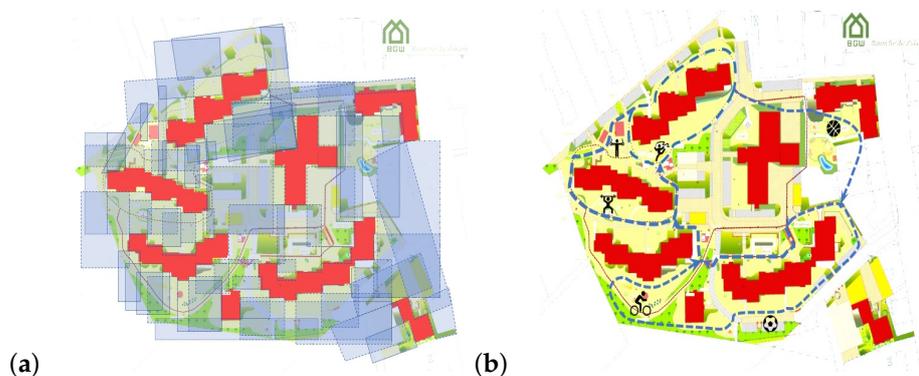
Cities can be considered to be among the largest and most complex artificial networks created by human beings. At the same time, the city is the biggest communication editor that determines the social and economic wellbeing not only of our present generation but also for those generations to come, as it provides an interface for our everyday mutual interactions. Sociologists think that isolation worsens an area's economic prospects by reducing opportunities for commerce and engenders a sense of isolation in inhabitants, both of which can fuel poverty and crime [19]. The speed and scale of urban growth require urgent global action to help cities prepare for growth and to avoid their being the future epicenters of poverty and human suffering. Physical segregation of minority groups dispersed over the spatially isolated pockets of streets and other inner-city areas being socially barricaded by railways and industries causes their economic marginalization [35].

Together with the “Neurocognition and Action – Biomechanics” research group of the Center for Cognitive Interaction Technology at Bielefeld University (Germany), we worked on a project aimed at improving the urban neighborhood of state-subsidized rental apartments located on Carlmeierstraße in Bielefeld.

Although it is often impossible to change the structure of built environments in urban neighborhoods, one can try to alter the social function of a place in order to increase its attendance and improve the local social climate in it. The idea was to encourage visiting the least-walkable paths and areas surrounding them by placing sports equipment in them. In order to quantify walkability in the urban neighborhood and identify the least-walkable paths through it, we used the approach discussed in Section 5. The spatial graph of the neighborhood was created by tessellation of the available walkable space, with overlapping

rectangles anchored at the built environment (see Figure 8a). Each rectangle shown in Figure 8a represented a vertex of the spatial graph, and each overlap of these rectangles was interpreted as an edge in the graph. These edges were weighted by the probabilities defined in (22), and then the least probable paths connecting the rectangles were identified with the use of Dijkstra's algorithm.

The university initiative received positive feedback from residents of the neighborhood.



**Figure 8.** The urban neighborhood of state-subsidized rental apartments located on Carlmeyerstraße in Bielefeld (Germany). (a) Tessellation of walkable space in the neighborhood with 66 overlapping rectangles anchored at the built environment. (b) The least-walkable paths in the neighborhood with places for sports equipment indicated.

## 7. Conclusions

In the present paper, we studied three important applications of CNE walks defined in finite connected graphs. The CNE walks define the properly normalized probability distributions for all nodes, subgraphs, and nodal sets of the graph at all time scales of the diffusion process.

The first application is related to the quantitative measures of *navigability* in graphs with respect to the different graph exploration strategies. In general, frequently visited graph nodes are predicted more efficiently than those that are little frequented. In particular, each node of the graph can be characterized by the navigability and strive potentials that depend on the node's function in the graph structure and the chosen graph exploration strategy. The navigability potential is defined as the probability that the node serves as a precursor for other nodes, while the strive potential of the node is interpreted as its neighborhood exploration time.

The second application is related to the mutual information quantifying the discernment of either of the diffusion modes by knowing the other, which we call *perspicacity*. We have shown that perfect perspicacity is possible only when both diffusion modes enjoy the same symmetry (grounded in the structural symmetry of the graph).

Finally, the third application is related to a *walkability* assessment of the graph, including identification of the most and least probable paths in that. We assume that the most walkable path in a graph is that most likely to be walked over a long enough observation time. The proposed walkability analysis was used for the assessment of walkability in urban neighborhoods and received positive feedback from the neighborhood residents.

The application of information theory methods to problems about graphs, in contrast to geometric, combinatoric, algorithmic, and algebraic approaches, can be called information graph theory. As it involves evaluating communication efficiency between individual systems' units at different time and connectivity scales, information graph theory is in demand for a wide range of applications, such as designing network-on-chip architecture and engineering urban morphology within the concept of the smart city.

**Funding:** This research received no external funding.

**Data Availability Statement:** The data are contained within the article.

**Acknowledgments:** The author is grateful to his institution for the administrative and technical support.

**Conflicts of Interest:** The author declares no conflict of interest.

### Abbreviations

The following abbreviations are used in this manuscript:

NoC	Network-on-chip
MCE	Microcanonical ensemble (of walks)
CNE	Canonical ensemble (of walks)
GCE	Grand canonical ensemble (of walks)

### References

1. Johnson, J. *Hypernetworks in the Science of Complex Systems*; World Scientific Series “Complexity Science”; 516 Sherfield building Imperial College London SW7 2AZ United Kingdom; Imperial College Press: London, UK, 2014; Volume 3.
2. Venkataraman, N.L.; Kumar, R. Design and analysis of application specific network on chip for reliable custom topology. *Comput. Netw.* **2019**, *158*, 69–76. [[CrossRef](#)]
3. Manzoor, M.; Mir, R.N. A Review Of Design Approaches For Enhancing The Performance of Nocs at Communication Centric Level. *Scalable Comput. Pract. Exp.* **2021**, *22*, 347–364. [[CrossRef](#)]
4. Alimi, I.A.; Patel, R.K.; Aboderin, O.; Abdalla, A.M.; Gbadamosi, R.A.; Muga, N.J.; Pinto, A.N.; Teixeira, A.L. Network-on-Chip Topologies: Potentials, Technical Challenges, Recent Advances and Research Direction. In *Network-on-Chip-Architecture, Optimization, and Design Explorations*; IntechOpen: London, UK, 2022.
5. Kaleem, M.; Isnin, I.F.B. A Survey on Network on Chip Routing Algorithms Criteria. In *Advances on Smart and Soft Computing. Advances in Intelligent Systems and Computing*; Saeed, F., Al-Hadhrani, T., Mohammed, F., Mohammed, E., Eds; Springer: Singapore, 2021; Volume 1188.
6. Gogoi, A.; Ghoshal, B.; Sachan, A.; Kumar, R.; Manna, K. Application driven routing for mesh based Network-on-Chip architectures. *Integr. Vlsi J.* **2022**, *84*, 26–36. [[CrossRef](#)]
7. Nassima, K.; Mouloud, K. A survey on fault-tolerant application mapping techniques for Network-on-Chip. *J. Syst. Archit.* **2019**, *92*, 39–52.
8. Romanov, A.; Myachin, N.; Sukhov, A. Fault-Tolerant Routing in Networks-on-Chip Using Self-Organizing Routing Algorithms. In Proceedings of the IECON 2021—47th Annual Conference of the IEEE Industrial Electronics Society, Toronto, ON, Canada, 13–16 October 2021; pp. 1–6.
9. Trajkovic, J.; Karimi, S.; Hangsan, S.; Zhang, W. Prediction Modeling for Application-Specific Communication Architecture Design of Optical NoC. *ACM Trans. Embed. Comput. Syst.* **2022**, *21*, 35. [[CrossRef](#)]
10. D’Acci, L.; Batty, M. (Eds.) *The Mathematics of Urban Morphology*. In *Birkhauser Series: Modeling and Simulation in Science*; Springer: Berlin/Heidelberg, Germany, 2020; ISBN 9783030123833.
11. Rassia, S.T.; Pardalos, P. (Eds.) *Smart City Networks*. In *Through the Internet of Things*; Springer Series Optimization and Its Applications; Springer: Cham, Switzerland, 2017; ISBN 978-3-319-61312-3.
12. Hillier, B.; Hanson, J. *The Social Logic of Space*; (1993, reprint, paperback edition ed.); Cambridge University Press: Cambridge, UK, 1984.
13. Hillier, B. *Space Is the Machine: A Configurational Theory of Architecture*; Cambridge University Press: Cambridge, UK, 1999; ISBN 0-521-64528-X.
14. Hillier, B. *The Common Language of Space: A Way of Looking at the Social, Economic and Environmental Functioning of Cities on a Common Basis*; Bartlett School of Graduate Studies: London, UK, 2004.
15. Vaughan, L.; Chatford, D.; Sahbaz, O. *Space and Exclusion: The Relationship between Physical Segregation. Economic Marginalization and Poverty in the City*; Paper presented to Fifth Intern; Space Syntax Symposium: Delft, The Netherlands, 2005.
16. Epstein, R.A.; Vass, L.K. Neural systems for landmark-based wayfinding in humans. *Philos. Trans. R. Soc. B Biol. Sci.* **2014**, *369*, 20120533. [[CrossRef](#)]
17. Macchi, O. The coincidence approach to stochastic point processes. *Adv. Appl. Probab.* **1975**, *7*, 83–122. [[CrossRef](#)]
18. Webler, F.; Andersen, M. Measurement in the Age of Information. *Information* **2022**, *13*, 111. [[CrossRef](#)]
19. Volchenkov, D. Grammar of Complexity: From Mathematics to a Sustainable World. In *Nonlinear Physical Science*; World Scientific Series; Higher Education Press: Beijing, China, 2018; pp. 1–3.
20. Volchenkov, D. Memories of the Future. Predictable and Unpredictable Information in Fractional Flipping a Biased Coin. *Entropy* **2019**, *21*, 807. [[CrossRef](#)] [[PubMed](#)]
21. Volchenkov, D. Infinite Ergodic Walks in Finite Connected Undirected Graphs. *Entropy* **2021**, *23*, 205. [[CrossRef](#)]
22. Volchenkov, D.; Suh, C.S. Statistical Mechanics of Long Walks in Dynamic Complex Networks: Statistical Arguments for Diversifying Selection. *Dynamics* **2022**, *2*, 252–269. [[CrossRef](#)]
23. Khinchin, A.I. *Mathematical Foundations of Information Theory*; Dover: New York, NY, USA, 1957.
24. Ramshaw, J.D. *The Statistical Foundations of Entropy*; World Scientific: Singapore, 2018.

25. Cover, T.M.; Thomas, J.A. *Elements of Information Theory*; Wiley: Hoboken, NJ, USA, 1991; p. 576.
26. Grabski, F. Discrete state space Markov processes. In *Semi-Markov Processes: Applications in System Reliability and Maintenance*; Elsevier: Amsterdam, The Netherlands, 2015; pp. 1–17, ISBN 978-0-12-800518-7.
27. Watanabe, S.; Accardi, L.; Freudenberg, W.; Ohya, M. (Eds.) *Algebraic Geometrical Method in Singular Statistical Estimation*; Series in Quantum Bio-Informatics; World Scientific: Singapore, 2008; pp. 325–336.
28. James, R.G.; Ellison, C.J.; Crutchfield, J.P. Anatomy of a bit: Information in a time series observation. *Chaos* **2011**, *21*, 037109. [[CrossRef](#)]
29. Travers, N.F.; Crutchfield, J.P. Infinite excess entropy processes with countable-state generators. *Entropy* **2014**, *16*, 1396–1413. [[CrossRef](#)]
30. Volchenkov, D. Assessing Pandemic Uncertainty on Conditions of Vaccination and Self-isolation. *Lobachevskii J. Math.* **2022**, *43*, 490–500. [[CrossRef](#)]
31. Minc, H. *Nonnegative Matrices*; John Wiley & Sons: New York, NY, USA, 1988.
32. Mowshowitz, A. Graphs, groups and matrices. *Proc. Summer Meet. Canad. Math. Congr. Congr. Numer.* **1971**, *4*, 509–522
33. Chan, A.; Godsil, C. Symmetry and Eigenvectors. In *Graph Symmetry, Algebraic Methods and Applications*; Hahn, G., Sabidussi, G., Eds.; Kluwer: Dordrecht, The Netherlands, 1997; pp. 75–106.
34. Godsil, C.; Royle, G. Graduate Texts in Mathematics. In *Algebraic Graph Theory*; Springer: New York, NY, USA, 2001; Volume 207.
35. Blanchard, P.; Volchenkov, D. Understanding Complex Systems. In *Mathematical Analysis of Urban Spatial Networks*; Springer Series; Springer: Berlin/Heidelberg, Germany, 2009; ISBN 978-3-540-87828-5.
36. Blanchard, P.; Volchenkov, D. Introduction to Random Walks and Diffusions on Graphs and Databases. In *Springer Series Synergetics*; Springer: Berlin/Heidelberg, Germany, 2011; Volume 10, ISBN 978-3-642-19591-4.
37. Adler, R.L.; Konheim, A.G.; McAndrew, M.H. Topological entropy. *Trans. Am. Math. Soc.* **1965**, *114*, 309–319. [[CrossRef](#)]
38. Delvenne, J.C.; Libert, A.S. Centrality measures and thermodynamic formalism for complex networks. *Phys. Rev. E* **2011**, *83*, 046117. [[CrossRef](#)]
39. Burda, Z.; Duda, J.; Luck, J.M.; Waclaw, B. Localization of the Maximal Entropy Random Walk. *Phys. Rev. Lett.* **2009**, *102*, 160602. [[CrossRef](#)]
40. Landau, E. Zur relativen Wertbemessung der Turnierresultate. *Dtsch. Wochensach* **1895**, *11*, 366–369.
41. Slater, J. The Theory of Complex Spectra. *Phys. Rev.* **1929**, *34*, 1293–1322. [[CrossRef](#)]
42. Vershik, A.M. *Asymptotic Combinatorics with Applications to Mathematical Physics a European Mathematical Summer School Held at the Euler Institute, St. Petersburg, Russia, July 9–20. 2001*; Springer: Berlin, Germany, 2003; p. 151, ISBN 978-3-540-44890-7.
43. Deng, N.; Zhou, W.; Haenggi, M. The Ginibre point process as a model for wireless networks with repulsion. *IEEE Trans. Wirel. Commun.* **2015**, *14*, 107–121. [[CrossRef](#)]
44. Miyoshi, N.; Shirai, T. A Cellular Network Model with Ginibre Configured Base Stations. *Adv. Appl. Probab.* **2016**, *46*, 832–845. [[CrossRef](#)]
45. Katz, N.M.; Sarnak, P. Zeroes of Zeta Functions and Symmetry. *Bull. AMS* **1999**, *36*, 1–26. [[CrossRef](#)]
46. Muir, T. Orthogonants. In *Treatise on the Theory of Determinants*; Revised and Enlarged by W. H. Metzler; Dover: New York, NY, USA, 1960; Chapter 14.
47. Golledge, R.G. (Ed.) Wayfinding Behavior. In *Cognitive Mapping and Other Spatial Processes*; John Hopkins University Press: Baltimore, MD, USA, 1999.
48. Woods, C.T.; Rudd, J.; Robertson, S.; Davids, K. Wayfinding: How Ecological Perspectives of Navigating Dynamic Environments Can Enrich Our Understanding of the Learner and the Learning Process in Sport. *Sport.-Med.-Open* **2020**, *6*, 51. [[CrossRef](#)]
49. Kohler, M.; Wehner, R. Idiosyncratic route-based memories in desert ants. *NEurobiology Learn. Mem.* **2004**, *83*, 1–12.
50. Jacobs, J. *The Death and Life of Great American Cities*; Knopf Doubleday Publishing Group: New York, NY, USA, 1992.
51. Dijkstra, E.W. A note on two problems in connexion with graphs. *Numerische Mathematik* **1959**, *1*, 269–271. [[CrossRef](#)]
52. Mehlhorn, K.; Sanders, P. Chapter 10. Shortest Paths. In *Algorithms and Data Structures: The Basic Toolbox*; Springer: Berlin/Heidelberg, Germany, 2008.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.