AIP Publishing

Random walk in random permutation set theory *⊙*

Jiefeng Zhou [®] ; Zhen Li [®] ; Yong Deng ■ [®]



Chaos 34, 093137 (2024)

https://doi.org/10.1063/5.0220154





Articles You May Be Interested In

Dempster-Shafer evidence theory-application approach

AIP Conf. Proc. (August 2023)

Characterizing unstructured data with the nearest neighbor permutation entropy

Modelling the eye movements of dyslexic children during reading as a continuous time random walk

Chaos (August 2023)







12 October 2024 05:41:18

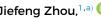
Random walk in random permutation set theory

Cite as: Chaos 34, 093137 (2024); doi: 10.1063/5.0220154 Submitted: 23 May 2024 · Accepted: 3 September 2024 · Published Online: 25 September 2024











Jiefeng Zhou, 1,a) D Zhen Li,2 D and Yong Deng Lb) D



AFFILIATIONS

- ¹Institute of Fundamental and Frontier Science, University of Electronic Science and Technology of China, Chengdu 610054, China
- ²China Mobile Information Technology Center, Beijing 100029, China
- ^{a)}Also at: Yingcai Honors College, University of Electronic Science and Technology of China, Chengdu 610054, China.
- ^{b)}Author to whom correspondence should be addressed: dengentropy@uestc.edu.cn and prof.deng@hotmail.com.

Also at: School of Medicine, Vanderbilt University, Nashville, Tennessee 37240, USA.

ABSTRACT

Random walk is an explainable approach for modeling natural processes at the molecular level. The random permutation set theory (RPST) serves as a framework for uncertainty reasoning, extending the applicability of Dempster-Shafer theory. Recent explorations indicate a promising link between RPST and random walk. In this study, we conduct an analysis and construct a random walk model based on the properties of RPST, with Monte Carlo simulations of such random walk. Our findings reveal that the random walk generated through RPST exhibits characteristics similar to those of a Gaussian random walk and can be transformed into a Wiener process through a specific limiting scaling procedure. This investigation establishes a novel connection between RPST and random walk theory, thereby not only expanding the applicability of RPST but also demonstrating the potential for combining the strengths of both approaches to improve problem-solving

Published under an exclusive license by AIP Publishing. https://doi.org/10.1063/5.0220154

Random walk models simulate natural processes by modeling step-by-step movements as random variables. The random permutation set theory (RPST) is an uncertainty reasoning framework extending evidence theory with ordered information. This study reveals a promising connection between RPST and random walks. By constructing a random walk model based on RPST properties and analyzing it through simulations, the key finding is that the RPST-generated random walk exhibits characteristics similar to a Gaussian random walk and can be transformed into a Wiener process (Brownian motion)—through scaling. Specifically, the RPST walk displays normally distributed step sizes, variance growing quadratically with steps, and convergence to Wiener processes in the limit. This novel link between RPST and stochastic processes implies possibilities for improved functionalities in various applications utilizing both methods.

I. INTRODUCTION

Random walk models have been used to simulate various natural processes. These models are particularly useful for understanding molecular-level dynamics (Ansari-Rad et al., 2012; Kessing et al., 2022; and Thompson et al., 2022), complex networks (Wang et al.,

2021), and so on. Correlated walks are specialized category of random walks. In correlated walks, the moving particles have a memory of their previous steps. This memory affects the direction of the next step, making the order of the steps important (Tojo and Argyrakis,

The Dempster-Shafer evidence theory (DSET), also known as evidence theory, is a framework used for reasoning under uncertainty (Dempster, 1967 and Shafer, 1976). In contrast to probability theory, DSET utilizes mass functions to assign beliefs to subsets of a power set, rather than to individual outcomes in the sample space, allowing for a more flexible approach to defining belief assignments. Moreover, DSET has been extended to complex evidence theory (Xiao et al., 2023 and Xiao and Pedrycz, 2023) and quantum evidence theory (Xiao, 2023a; 2023b), which are applied to various fields, such as time series analysis (Cui et al., 2022; Qiang et al., 2022; Contreras-Reyes and Kharazmi, 2023; Kharazmi and Contreras-Reyes, 2024; and Zhang and Xiao, 2024), quantum Dempster's rule of combination (He and Xiao, 2024), software risk assessment (Chen and Deng, 2024), and so on. However, DSET struggles to handle ordered information in certain real-world problems. To address this limitation, the random permutation set theory (RPST) is proposed (Deng, 2022). By introducing the concept of permutation events, RPST effectively considers the order of elements and expands the power set and mass function into the permutation event space (PES) and permutation mass function (PMF). Similar to Shannon entropy (Shannon, 1948) in probability and Deng entropy (Zhao *et al.*, 2024) in DSET, the RPS entropy is proposed by Chen and Deng (2023) to quantify the uncertainty in RPST. The orthogonal sum in DSET is further extended in RPST by considering the order of the elements in evidence (Wang *et al.*, 2024).

One research area of focus in DSET and RPST is its physical implications. For instance, in DSET, Li and Xiao (2023) derived normal distribution from maximum Deng entropy, and Zhao et al. (2024) found intriguing linearity in Deng entropy. In RPST, Zhan et al. (2024) established a bridge between thermodynamic entropy and information entropy from the perspective of measurement, delved into the PMF condition for maximum entropy in RPST. Zhao et al. (2023) further demonstrated that the information dimension associated with this PMF condition is 2, similar to the fractal dimension of Brownian motion. Moreover, the mean square distance in a Brownian motion denoted as $\bar{r^2}$ is proportional to the time elapsed. While in our previous research, it is showed that the limit form of maximum RPS entropy is $e \cdot (n!)^2$, exhibiting similarities to \bar{r}^2 (Zhou et al., 2024). These collective discoveries hint at a potential connection between RPST and Brownian motion, or random walk in mathematics.

In this paper, we conduct an in-depth analysis of RPST and construct random variables based on its properties. We then generate a random walk using these random variables. Finally, we demonstrate that this type of random walk shares similarities with a Gaussian random walk and can be converted into a Wiener process through scaling.

In general, we builds a bridge between RPST and random walk theory, revealing their potential applications in modeling complex real-world phenomena. For example, in fields such as financial modeling, where asset prices exhibit random walk behavior, the introduction of RPST could enhance the accuracy of predictions by incorporating the order of events. Similarly, in molecular dynamics, where the paths of particles are critical, the connection between RPST and random walks could lead to better simulations of molecular behavior. This connection not only broadens the utility of RPST but also enhances the problem-solving capabilities of random walk models, particularly in fields requiring precise handling of ordered or sequential data. By integrating the strengths of both methodologies, researchers can develop more robust models that are better suited to capture the complexity of these systems.

The article is structured as follows. Section II introduces some key concepts related to this work. In Sec. III, the construction of RPST-generated random walk is presented. Finally, this article is summarized in Sec. IV. The proof of this work is attached in Appendix A.

II. PRELIMINARIES

Some key concepts about this work are introduced in this section.

A. Sample space and mass function

Definition 1 (Sample space). A sample space Ω is a mathematical set that contains all possible base events E_i , and the

cardinality of sample space is denoted as $|\cdot|$. The *power set* of Ω is marked as 2^{Ω} .

$$\Omega = \{E_1, E_2, E_3, \dots, E_n\}, \quad |\Omega| = n.$$
 (1)

Definition 2 (Mass function). A mass function $\mathcal{M}(\cdot)$ is a function that assigns a belief to each subset of a sample space Ω , $\mathcal{M}: 2^{\Omega} \to [0,1]$, with the following constraints:

$$\sum_{i} \mathcal{M}(i) = 1, \quad \mathcal{M}(i) \ge 0, \quad \mathcal{M}(\emptyset) = 0.$$
 (2)

B. Random permutation set theory

By introducing ordered information, the random permutation set theory (RPST) successfully extends the scope of evidence theory. Some fundamental concepts of RPST are introduced below.

1. Random permutation set

Definition 3 (Permutation event space, PES). The permutation event space (PES) is a set that contains all possible permutations of base events of Ω ,

$$PES(\Omega) = \{ p_{ij} | i = 0, 1, \dots, n; j = 1, 2, \dots, P(n, i) \}$$

$$= \{ \emptyset, [E_1], [E_2], \dots, [E_n], [E_1, E_2], [E_2, E_1], \dots,$$

$$[E_{n-1}, E_n], [E_n, E_{n-1}], \dots,$$

$$[E_1, E_2, \dots, E_n], [E_n, E_{n-1}, \dots, E_1] \},$$
(3)

where P(n, i) = n!/(n - i)! is the *i*-permutation of *n*.

Definition 4 (Permutation mass function, PMF). A permutation mass function (PMF) M is a mapping $M: PES(\Omega) \rightarrow [0,1]$, with constraints

$$M(\emptyset) = 0, \quad \sum_{p \in PES(\Omega)} M(p) = 1.$$
 (4)

The **random permutation set (RPS)** consists of a permutation event from $PES(\Omega)$ and its associated permutation mass function (PMF) $M: RPS(\Omega) = \{A, M(A) | A \in PES(\Omega)\}.$

2. RPS entropy

Similar to entropy methods as uncertainty measure in evidence theory, RPS entropy has been proposed recently (Chen and Deng, 2023). What is more, the maximum RPS entropy and its limit form are also introduced and proved (Zhan et al., 2024 and Zhou et al., 2024).

Definition 5 (RPS entropy). The RPS entropy of a RPS $RPS(\Omega) = \{A, M(A)\}$ is defined as

$$H_R PS(M) = -\sum_{A \in PES(\Omega)} M(A) \log (M(A)/(F(|A|) - 1),$$
 (5)

where |A| is the cardinality of permutation event A, and $F(i) = \sum_{i=0}^{i} P(i,j)$.

RPS entropy is fully compatible with Deng entropy (Zhao et al., 2024) as used in evidence theory and Shannon entropy (Shannon, 1948) in probability theory. Such uncertainty measures have been provided insights for real-world applications

(Huang et al., 2023a; 2023b; Xiao et al., 2023; and Deng et al., 2024) and other kinds of uncertainty measures like information dimension (Zhao et al., 2023 and Ortiz-Vilchis et al., 2024), information distance (Xiao and Pedrycz, 2023), negation (Kharazmi and Contreras-Reyes, 2023 and Kharazmi et al., 2023), and so on (Kharazmi et al., 2023) and Kharazmi and Contreras-Reyes, 2024).

Zhan et al. (2024) derived and proved the following PMF condition of maximum RPS entropy:

$$M(A) = \frac{F(|A|) - 1}{\sum_{i=1}^{n} [P(n, i)(F(i) - 1)]}.$$
 (6)

The corresponding maximum RPS entropy for such PMF condition is then expressed as

$$H_{max-RPS} = \log \left(\sum_{i=1}^{n} [P(n,i)(F(i)-1)] \right).$$
 (7)

The limit form of maximum RPS entropy can be simplified as (Zhou et al., 2024)

$$H_{max-RPS} \approx e \cdot (n!)^2$$
. (8)

This elegant result offers valuable insights into the physical significance of RPST. In a study of Brownian motion, Einstein (1956) demonstrated that the mean square displacement is directly proportional to the elapsed time, expressed as $\bar{r^2} \propto t$. This prompts us to investigate a potential relationship between $(n!)^2$ and $\bar{r^2}$. Furthermore, Zhao *et al.* (2023) identified that the information dimension of the PMF associated with maximum RPS entropy is 2, which aligns with the fractal dimension of Brownian motion. Collectively, these findings suggest a possible link between RPST and Brownian motion, or random walk theory.

C. Random walk

How to model and measure the uncertainty and dynamics of the system Wang *et al.* (2017; 2018; 2020; 2022) have attracted much attention. Random walk is a fundamental topic in probability theory. It is a type of stochastic process, which is a sequence of random variables that evolve over time. Random walk is formed by the successive summation of independent and identically distributed (i.i.d.) random variables (Lawler and Limic, 2010).

1. General random walk

Definition 6 (General random walk). For $\forall t \in \mathbb{N}^+ = \{1, 2, 3 \dots \}$, let $S_t \in \mathbb{R}^d$, given a proper probability distribution $P : \mathbb{R}^d \to (0, 1]$ and a group of i.i.d. random variables $\{X_t | X_t \in \mathbb{R}^d, t \in \mathbb{N}^+ \}$, the general random walk S_n with step size distribution P can be considered the time-homogeneous Markov chain, defined by a summation of $\{X_t\}$,

$$S_n = S_0 + X_1 + X_2 + \ldots + X_n, \tag{9}$$

where $S_0 \in \mathbb{R}^d$ is the starting point.

One well-studied variant is the Gaussian random walk, which has a step size distribution of normal distribution $N(0, \sigma^2)$.

2. Wiener process

The Wiener process, also known as Brownian motion, is a fundamental concept in probability theory and stochastic processes. It represents the limiting behavior of a one-dimensional random walk as the step size approaches zero and the number of steps approaches infinity.

Definition 7 (Wiener process). For $\forall n \in \mathbb{Z}_+$, W(t) is a Wiener process if

$$W(t) = \lim_{n \to \infty} W_n(t) = \frac{1}{\sqrt{n}} \sum_{1 \le i < |nt|} \xi_i, t \in [0, 1], \xi \sim N(0, 1). \quad (10)$$

A Wiener process $\{W(t), t > 0\}$ has the following properties:

• W(0) = 0;

15 return (V_x, V_y)

- for 0 < s < t, the increments $W(t) W(s) \sim N(0, t s)$;
- For any non-overlapping interval $[s_i, t_i]$, the group of random variables W(t) W(s) are independent of each other;
- W(t) is almost surely continuous in t.

Those properties will be analyzed in our proposed stochastic process.

ALGORITHM 1. Random Variable Generator.

```
Result: A vector in perpendicular coordinates (V_x, V_y)
            representing the addition of component vectors.
  Input: Integer n indicating the length of the set.
  Output: Vector in perpendicular coordinates (V_x, V_y).
1 M \leftarrow [0]_{(n!,n)}
   /st Initialize a zero matrix M with dimensions
       (n!,n).
2 S ← {1,2,...,n}
  /* Initialize the set S = \{1, 2, \dots, n\}
3 for each permutation list p_i in all permutations of S. do
4 M_i \leftarrow p_i
6 p_s \leftarrow \texttt{RandomChoice}(M)
   /* Select a possible permutation sequence p_s
       from matrix M evenly based on uniform
7 V_x, V_y \leftarrow 0
   /* Initialize sum of x, y components vectors.
s for i \leftarrow 1 to n do
      \theta_i \leftarrow \frac{2\pi}{n} \cdot i
      x_i \leftarrow a_i \cdot \cos(\theta_i)
    \begin{vmatrix} y_i \leftarrow a_i \cdot \sin(\theta_i) \\ V_x \leftarrow V_x + x_i \\ V_y \leftarrow V_y + y_i \end{vmatrix}
```

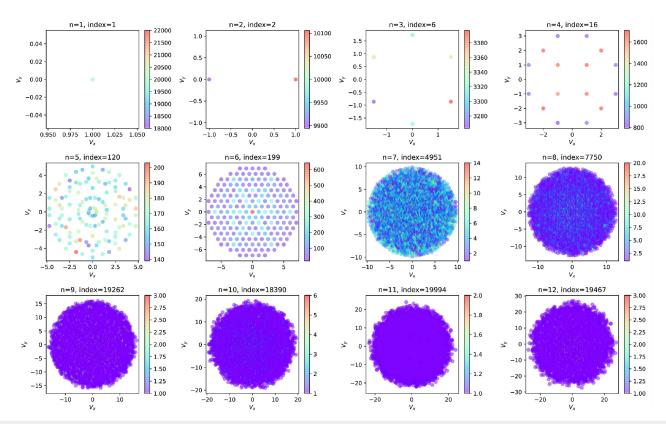


FIG. 1. Visualization of random variables generation with n ranging from 1 to 12 in 20 000 simulations. Each point on the graph represents a possible random variable, with the color indicating the frequency of occurrence.

III. EXPLORE RANDOM WALK IN RANDOM PERMUTATION SET

In this section, we will give an in-depth exploration of random walk in random permutation set theory. Now let us review the motivation discussed earlier: the RPST brings order information of events to expand evidence theory, and the order information can be viewed as time sequence information since time has fixed order and flows in one direction. The term "random" inspired us to find the connection between RPST and random walks in stochastic process.

When generating a random walk based on RPST, it is important to consider the order information present in RPST. For convenience, we use list $[a_1, a_2, \ldots, a_n]$ to express order information and simulate two-dimensional random walk. First, the random variable should be defined.

One situation where the ordered information is important is the matrix multiplication, because matrix multiplication does not hold commutative property, i.e., AB = BA for most of the matrix A, B. This inspired us to use matrix to generate a random variable.

Given a permutation sequence, $PerS_{n\times 1} = (a_1, a_2, \dots, a_n)^{\top}$ and a arbitrary vector $\vec{V}_0 = (x, y)^{\top}$, we want to output a random variable vector $\vec{V}_i = (V_x, V_y)^{\top}$ top. This can be done by the following computation. First, we randomly generated some invertible matrices M_N

= $(M_1, M_2, ..., M_i, ..., M_n)^{\top}$, then we compute $Vec_i = a_i \cdot M_i \vec{V}_0$ for each M_i and each a_i , getting n component vectors Vec_i . Then, we have a summation vector $\vec{V}_i = \sum_i Vec_i$ by adding all component vectors

For the convenience of illustration, we use two-dimensional rotation matrix $R(\theta)_{2\times 2}$ to replace M_i in the following way:

$$R(\theta)_{2\times 2} = \begin{pmatrix} \cos\theta & \sin\theta\\ \sin\theta & \cos\theta \end{pmatrix},\tag{11}$$

$$M_N = \left(R(\theta_0), R^2(\theta_0), \dots, R^i(\theta_0), \dots, R^n(\theta_0) \right)^\top. \tag{12}$$

We use the following algorithm to generate a random variable.

A. Generating random variables

Definition 8 (Random Variable Generator, RVG). Given a positive integer *n*, the random variable generator (RVG) is defined by Algorithm 1.

Algorithm 1 takes an integer n as input, outputting a vector in perpendicular coordinates marked as a random variable. The n denotes the number of component vectors, and the cardinality of a possible permutation sequence $p_i = (a_1, a_2, \ldots, a_n)$. The reason of

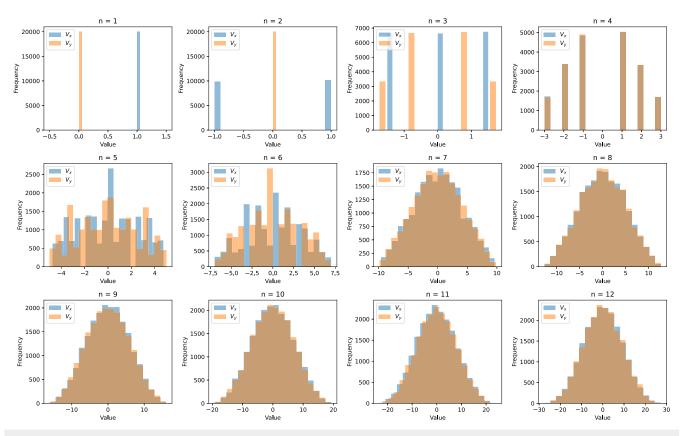


FIG. 2. Histogram of random variables with respect to V_x and V_y .

choosing possible permutation sequence is discussed in Sec. III B. After the possible permutation sequence p_i is selected, the numbers in it indicate the length of each component vector. As for the direction of each component vector, we choose to divide 2π into n piece evenly, so each component vector \vec{v}_j can be defined as $\vec{v}_j = (a_j, 2 \cdot j \cdot \pi/n)$. Then, we output the random variable $(V_x, V_y) = \vec{V}_i$ by adding all component vectors.

As shown in Fig. 1, we simulate 20 000 random variables with n ranging from n=1 to n=12. The index at the top of each sub-figure denotes the number of possible random variables (when $n \geq 7$, the number may be inaccurate due to limited simulation), while the color in each node represents the frequency in simulation. All numerical results are rounded to eight decimal places.

For n = 1, 2, 3, there are 1!, 2!, 3! kinds of values of random variables, respectively. While for n = 4, there are not 4! = 24 but 16 different values of random variable, as shown in the figure. This can be predicted, because when n = 4, each component vector has a fixed direction, which are $\pi/2$, π , $3\pi/2$, and 2π , respectively. This means each of these vectors points either horizontally or vertically. So each sum in the resulting vector's x or y direction can be produced in four ways: [1(2-1, 3-2, 4-3), 2(3-1, 4-2), 3(4-1)], which yields four different combinations: (1,3), (3,1), (2,2), (1,1). Since each of four ways

implies a rotation direction, which in turn leads to the x and y coordinates of the vector being multiplied by either +1 or -1. Thus, there are $4 \times 4 = 16$ unique possibilities.

This explanation can be extended to the cases of n = 5 and n = 6. However, the number of possible random variables grows rapidly when $n \ge 7$, compared with the simulation of n = 6. This is intuitive due to the rapid growth of factorial.

We examined the specifics of each random variable simulation concerning V_x and V_y . The histogram in Fig. 2 displays the distribution of V_x , V_y values in 20 000 simulations. The x-axis and y-axis in each sub-figure represent the value and frequency, respectively. The symmetrical distribution of frequencies in each interval suggests that the expected values of V_x , V_y should be zero, a finding supported by the results in Fig. 3. It is also anticipated that as the number of simulations denoted by n tends toward infinity, V_x , V_y will converge to a normal distribution.

Another important statistic property is the variance, Fig. 3 shows the variance of V_x , V_y in different values of n. As n increases, the variance of both V_x and V_y will grow like binomial function, which means $Var(V_{x,y}) \propto (n^2 + n)$. This variance growing speed property is another necessary feature of random walk. In Fig. 3, we compared the variance of both V_x and V_y , the linearity between them indicates that V_x and V_y are independent and symmetrical,

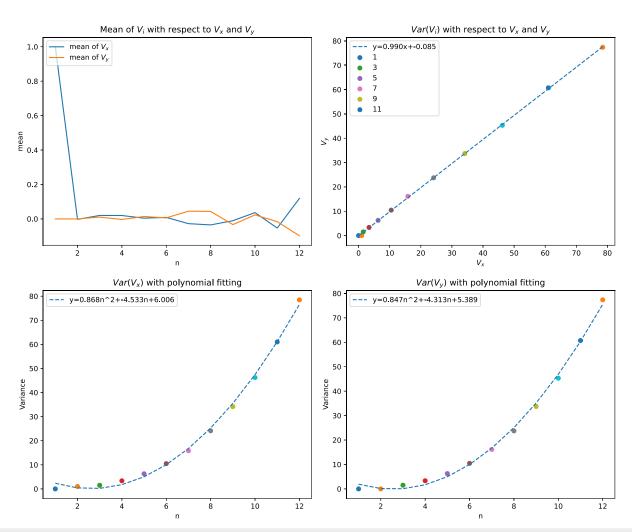


FIG. 3. Variance and mean value of V_i with respect to V_x and V_y .

ensuring that this simulation method is like Wiener process, which is invariant to rotations.

B. Simulating random variables for random walk

One common way to simulate random walk is adding a sequence of i.i.d. random variables. For example, \vec{V}_i form a normal distribution $N(\mu, \sigma^2)$, where μ and σ are the mean and standard deviations of the normal distribution, respectively. Then, the sum of normally distributed random variables is a random walk (Lawler and Limic, 2010),

$$S_t = \sum_{i=0}^t \vec{V}_i,\tag{13}$$

where \vec{V}_i is marked as a step, V_0 is the starting value of the random walk, and t is the number of steps. Inspired by such method, we tend

to use such method with random variables simulated from RPST to generate random walk.

To generate a ideal random variable \vec{V}_i , we first should determine the length of the possible permutation sequence, i.e., n in RVG.

Based on the maximum RPS entropy, an natural idea is to delve a distribution from RPST and use it as probabilities associated with each possible permutation sequence. Given a fixed set $\Lambda = \{\lambda_1, \lambda_2, \dots, \lambda_n\}$, the belief assigned to possible permutation sequence whose cardinality is identical is the same. When the length of a possible permutation sequence is determined, then we can select one of the possible permutation sequence evenly as our probabilities association method, and that is why we use uniform distribution as the probabilities associated with each possible permutation sequence in Sec. III A.

Definition 9 (RPST distribution). Given a maximum length of permutation sequences N, there are P(N, n) choices to select a

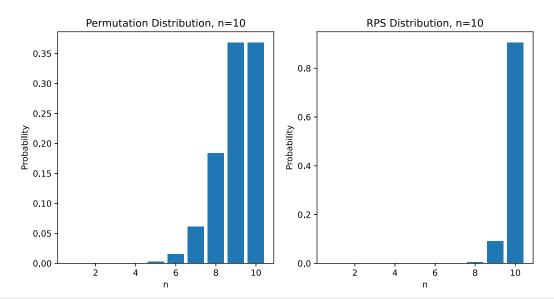


FIG. 4. Discrete probability distribution of $P_{Per}(n|N)$ and $P_{RPS}(n|N)$ with N=10.

possible permutation sequence with a length of n. The possibility of selecting n as the length of possible permutation sequence, when combined with the maximum RPS entropy, defines the RPST distribution as follows:

$$P_{RPS}(n|N) = P(N,n) \cdot \mathcal{M}_{i=n,j} = \frac{P(N,n)[F(n)-1]}{\sum_{i=1}^{N} [P(N,i)(F(i)-1)]}.$$
 (14)

To illustrate the validity of the proposed method, we consider the following way to select a possible length n for permutation sequence with the same probability as follows.

Definition 10 (Permutation distribution). Given a maximum length of permutation sequences N, there are $\sum_{i=1}^{N} i!$ kinds of permutation sequences, the permutation distribution is defined to choose a possible length n based on the number of permutation cases.

$$P_{Per}(n|N) = \frac{P(N,n)}{\sum_{i=1}^{N} P(N,n)} = \frac{P(N,n)}{F(N)-1} = \frac{P(N,n)}{[e \cdot N]},$$
 (15)

i.e., the possibility of selecting a possible sequence length n is in proportion to the magnitude of permutation P(N, n). In other words, given a maximum length of permutation sequences N, the

TABLE I. The last six elements' probability assignment of distribution $P_{Per}(n|N)$ and $P_{RPS}(n|N)$, all results are rounded to five digits.

	N	n = N - 5	n = N - 4	n = N - 3	n = N - 2	n = N - 1	n = N	$\sum_{n=N-5}^{N} P(n N)$
$P_{Per}(n N)$	6	3.0700×10^{-3}	1.5340×10^{-2}	6.1350×10^{-2}	1.8405×10^{-1}	3.6810×10^{-1}	3.6810×10^{-1}	1.0000×10^{-0}
	10	3.0700×10^{-3}	1.5330×10^{-2}	6.1310×10^{-2}	1.8394×10^{-1}	3.6788×10^{-1}	3.6788×10^{-1}	9.9941×10^{-1}
	: 18	3.0700×10^{-3}	1.5330×10^{-2}	6.1310×10^{-2}	1.8394×10^{-1}	3.6788×10^{-1}	3.6788×10^{-1}	9.9941×10^{-1}
	:							
	∞	$\frac{1}{5!e}$	$\frac{1}{4!e}$	$\frac{1}{3!e}$	$\frac{1}{2!e}$	$\frac{1}{e}$	$\frac{1}{e}$	$\frac{163}{60e}$
$P_{RPS}(n N)$	6	0.0000×10^{-0}	7.0000×10^{-5}	1.0800×10^{-3}	1.3820×10^{-2}	1.4035×10^{-1}	8.4468×10^{-1}	1.0000×10^{-0}
	10	0.0000×10^{-0}	1.0000×10^{-5}	2.1000×10^{-4}	5.0200×10^{-3}	9.0430×10^{-2}	9.0433×10^{-1}	1.0000×10^{-0}
	: 18	0.0000×10^{-0}	0.0000×10^{-0}	3.0000×10^{-5}	1.5500×10^{-3}	5.2550×10^{-2}	9.4587×10^{-1}	1.0000×10^{-0}
	:							
	∞	0	0	0	0	0	1	1

elements' probability assignment for those two distribution. Based

ALGORITHM 2. Random walk generator.

Result: A matrix $T_{1\times 2}$ representing the discrete time stochastic process.

Input: An integer t indicating the number of time steps, a selection method $P(P_{Per}(n|N))$ and $P_{RPS}(n|N))$, the maximum sequence length N.

Output: A matrix T with size of $t \times 2$.

- 1 $Len_Set \leftarrow GenLen(P, N)$
 - /* Create a set of lengths with constraints to maximum length of N, the set is used for generating random variables.
- **2 for** i = 2 **to** t **do**
- $Temp_Vec = RVG(Len_Set_{i-1}) T_i \leftarrow T_{i-1} + Temp_Vec$
- 4 end
 - /* Generate value of random walk at each time
 step.
- 5 return T

probability of selecting a possible permutation sequence from all $|e \cdot n!|$ sequences is $1/(|e \cdot n!|)$.

We plot the discrete probability distribution of $P_{Per}(n|N)$ and $P_{RPS}(n|N)$ with N=10 in Fig. 4. Table I lists the details of the last six

C. Generating random walk with random variables

Using Sec. III B as a construction of generating random walk, we design the following algorithm to generate random walk with random variables.

Definition 11 (Random walk generator, RWG). Given a positive integer T denoted as time steps, the maximum length of permutation sequence N, and the distribution method P(n|N), the random walk generator (RWG) is defined by Algorithm 2.

Algorithm 2 takes the number of time steps t, a distribution P used for selection method, and N indicating the maximum length of permutation sequence, as inputs, returning a matrix T storing values at each time step.

Figure 5 shows results across different N and P. The color map illustrates the temporal evolution of the random walk's trajectory. As N increases, the discrete-time stochastic process T, which

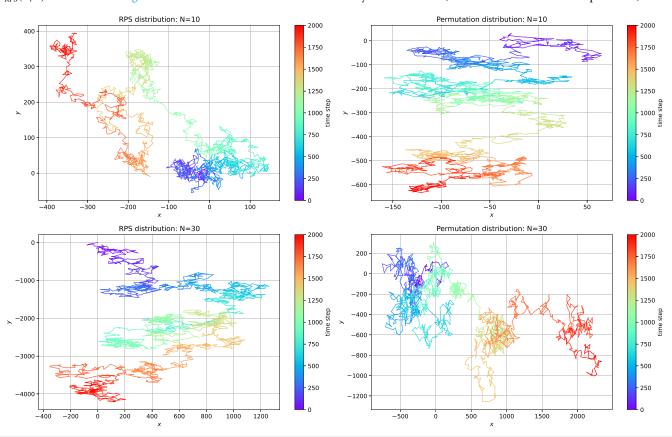


FIG. 5. Visualization of random walk from distribution $P_{Per}(n|N)$ and $P_{RPS}(n|N)$, where color map is showing the time steps.

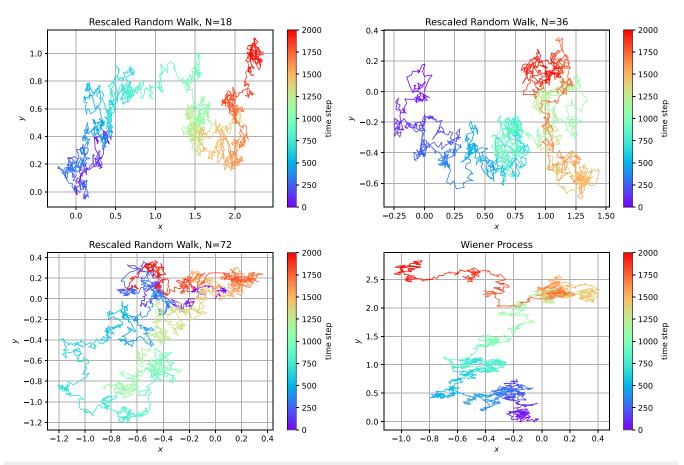


FIG. 6. Scaled random walk with different N and the Wiener process with a time step of 2000 and variance control factor $\varrho=24$.

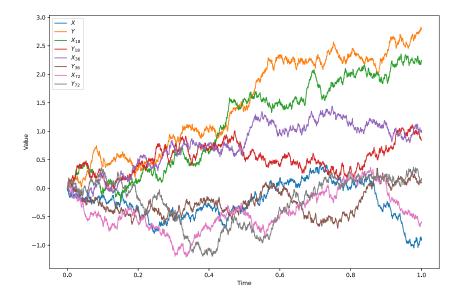


FIG. 7. Component value of random walk and the Wiener process of Fig. 6, the 2000 time steps are converted to time $t \in [0, 1]$ for convenience and ϱ is set to 24 $\varrho = 24$.

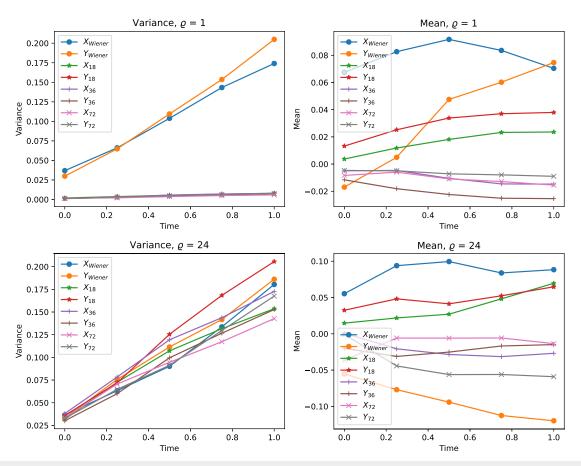


FIG. 8. Variance (left) and mean (right) value of the Wiener process W(t) and limit scale form of random walk from RPST $RW_{n,N}(t)$ across various N and variance control factor, time steps, and the number of simulations are set to 100, 200, respectively.

is generated from RPST distribution P_{RPS} , shows a motion pattern resembling random walk, characterized by randomly distributed points in space.

Comparing the results of N = 10 and N = 30, the motion exhibits stochastic self-similarity as in random walk. This is because at each time step, this RPST-generated motion will walk through the space for each n directions (n being the possible length of permutation sequence with a maximum length of N). These n paths can be decomposed into x and y directions in perpendicular coordinates, similar to the two-dimensional random walk where the walker randomly chooses one of two perpendicular directions with a fixed step size.

To compare the proposed method's limit scale form with the Wiener process, we employ a method similar to Definition 7, to simulate the limit scale form of the RPST-generated random walk,

$$RW_{n,N}(t) = \frac{\sqrt{\varrho}}{N\sqrt{N}} \frac{1}{\sqrt{n}} \sum_{1 \le i \le \lfloor nt \rfloor} \vec{V}_i, t \in [0,1], \tag{16}$$

where N is the maximum length of a permutation sequence, n is the number of time step, and ϱ is a variance control factor that scales

the variance of RPST-generated random walk . As $N, n \rightarrow infty$, $RW_{n,N}(t)$ toward to a Wiener process, the details are discussed in Appendix A.

The only difference to the Wiener process as a limit scale form of random walk is that the re-scaling factor $\sqrt{\varrho}/(N\sqrt{N})$. This is due to the fact that the random variables generated from RPST have variance growing like binomial function, as shown in Fig. 3. Thus, this redesigned re-scaling factor ensures the variance of random variables is invariant to n.

Since simulations are discrete, and for convenience of illustration, we generate the proposed stochastic process T with a time step of 2000 and re-scale to $RW_{n,N}(t)$ with setting $\varrho=24$.

As shown in Fig. 6, the scaling RPST-generated random walk do visually seem the same as standard Wiener process, not only the randomly walked point path but also the boundaries. More details about the component values about *X* and *Y* axis are plotted in Fig. 7. Based on this result, it seems that the proposed stochastic process converges to the Wiener process as *N* increases. However, additional verification is required before reaching a definitive conclusion.

In Fig. 8, the mean and variance values of various stochastic processes are compared to the Wiener process across

different ϱ values, with time steps and sample processes limited to 100 and 200, respectively. The five intervals are set to [0, 20), [0, 40), [0, 60), [0, 80), [0, 100) to minimize errors.

Results show that all proposed methods with different N exhibit properties similar to the Wiener process in terms of mean value and variance, where the mean value is zero and variance scales with time steps. Compared the sub-figures in Fig. 8, the difference lies in the slope of variance, which is why we introduce the variance control factor ϱ to regulate the variance of the proposed stochastic process. This ensures that the variance of the process aligns with that of a Wiener process.

From the results and analysis presented above, it is evident that as the sample size (*N*) increases, the RPST-generated random walk converges to a Wiener process, which is the limit scale form of a two-dimensional random walk. This demonstrates the successful derivation of a random walk from RPST.

IV. CONCLUSION

Random permutation set theory (RPST) is a promising extension of evidence theory that introduces ordered information to its reasoning framework. The indexed order in RPST can be viewed as a time series, which motivates the exploration of a connection between RPST and random walk, a fundamental topic in probability theory. This paper demonstrates that RPST can be used to construct a Gaussian random walk and, in the limit, a Wiener process. The established link between RPST and random walk provides insights into the physical meaning of RPST and enables its application in existing random walk domains. This not only expands the application scope of RPST but also provides insights for combination the strengths of both RPST and random walk for problem-solving.

Future investigations should concentrate on overcoming the limitations of current study. This may involve elucidating the physical implications of RPST through its association with random walks. Subsequently, the application of this random walk model to realworld scenarios, such as epidemiological modeling, financial market analysis, and machine learning algorithm, could be explored.

ACKNOWLEDGMENTS

The work is partially supported by National Natural Science Foundation of China (Grant No. 62373078).

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Jiefeng Zhou: Conceptualization (equal); Formal analysis (equal); Investigation (equal); Writing – original draft (equal); Writing – review & editing (equal). **Zhen Li:** Validation (equal). **Yong Deng:** Funding acquisition (equal); Project administration (equal); Supervision (equal); Writing – review & editing (equal).

DATA AVAILABILITY

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

APPENDIX: PROOF OF DERIVING RANDOM WALK FROM RPST

In this section, we will analyze the RPST-generated random walk in detail and demonstrate its similarities with random walk in mathematics.

1. Analysis on random variables

In Sec. III, the random variables are first defined for generating random walk. these variables are generated using the order property of RPST. As a simulation method, its important statistic properties like expected value and variance should be reviewed.

a. Expected value analysis. **Lemma 1.** (Expected value of a random variable). The expected value of a random variable generated with RVG is zero, namely,

$$\mathbb{E}\left[V_i\right] = 0. \tag{A1}$$

Proof. As described in Algorithm 1, when dealing with an integer set of length N, the likelihood of selecting a specific permutation sequence is equal, with a probability of

$$P\{V = V_i\} = \frac{1}{N!}.$$
 (A2)

To determine the expected value in each direction, we calculate the frequency of numbers appearing in a fixed direction, such as $\frac{2\pi \cdot i}{N}$. The magnitude of this direction in a simulation is determined by

$$|V_{component}| = P\{V = V_i\} \cdot \sum_{i=1}^{N} [j \cdot (N-1)!] = \frac{1+N}{2}.$$
 (A3)

Due to symmetry in direction generation and identical magnitudes, the resultant sum vector (V_x, V_y) is anticipated to yield a value of 0. Thus, the expected value of (V_x, V_y) or V_i is

$$\mathbb{E}\left[\left(V_{x}, V_{y}\right)\right] = \mathbb{E}\left[V_{i}\right] = 0. \tag{A4}$$

Figure 3 also displays the mean value of V_x , V_y in 20 000 simulations, suggesting the expected value of \vec{V}_i is zero.

b. Variance analysis. Variance is a measure of dispersion, which is pretty useful in generating random walk. As shown in Fig. 3, the variance of V_x and V_y are quantitatively identical, and both of them exhibit a binomial growth rate with respect to N.

Lemma 2 (Variance of a random variable). The variance of a random variable generated from RVG is a binomial function on N, namely:

$$Var(\vec{V}_i) \propto (N^2 + N).$$
 (A5)

Proof. In Fig. 3, it can be directly observed that the variance of both V_x and V_y is in proportion to N^2 . This relationship can be explained from the following perspectives.

As indicated in Appendix 1 a, the expected value of magnitude in each direction is proportional to the maximum length of permutation sequence N, while the expected value of \vec{V}_i remains zero, as demonstrated in Lemma 1. This ensures that as N increases, the distribution of random variables maintains its symmetry, resembling a round boundary. The size of this boundary is determined by the value of N, as shown in Fig. 1. Therefore, there exists a critical value N_0 , such that when $N_2 > N_1 \ge N_0$, random variables \vec{V}_{i2} generated with $N = N_2$ can be represented by the random variables \vec{V}_{i1} generated with $N = N_1$, denoted as

$$\vec{V}_{i2} = f(N_2 - N_1) \cdot \vec{V}_{i1},\tag{A6}$$

where f(x) is a function $\mathbb{R} \to \mathbb{R}$, as shown in Appendix 1 a. Then based on the propagation property of variance:

$$Var(aV) = a^2 Var(V), \tag{A7}$$

where a is a constant, and Appendix 1 b, one can easily delve such result in Lemma 2.

2. Analysis on permutation distribution and RPST distribution

Lemma 3 (Limit form of RPST distribution). When $N \rightarrow infty$, the RPST distribution will converge to the following form:

$$\lim_{N \to \infty} P_{RPS}(n = N|N) = 1 \tag{A8}$$

Proof. The RPST distribution is based on the maximum RPS entropy, this distribution will surely converge to P(n = N|N) = 1, as suggested in Table I. This result is determined by its definition on Eq. (14). In our previous work (Zhou *et al.*, 2024), we proved that

$$\lim_{N \to \infty} \sum_{i=1}^{N} \left[P(N, i) \left(F(i) - 1 \right) \right] - e \cdot (N!)^2 = 0, \tag{A9}$$

$$F(N) - 1 = \lfloor e \cdot N! \rfloor - 1 \tag{A10}$$

Compared with P(N, n)[F(n) - 1], we get

$$\lim_{N \to \infty} P_{RPS}(n = N|N) = \frac{N! \left(\lfloor e \cdot N! \rfloor - 1 \right)}{e(N!)^2}$$

$$= \lim_{N \to \infty} \frac{N! \lfloor e \cdot N! \rfloor}{e(N!)^2} - \lim_{N \to \infty} \frac{1}{e(N!)^2}$$

$$= 1 - 0$$

$$= 1 \tag{A11}$$

This result ensures that when N is bigger enough, this distribution will converge to the following probability distribution:

$$P_{RPS}(n|N) = \begin{cases} 1, & n = N; \\ 0, & others. \end{cases}$$
 (A12)

This probability distribution can be explained by the maximum entropy principle. This principle states that the distribution with the highest entropy is the most likely to represent the current state of a system. Therefore, the larger the value of N, the more likely it is that the system will choose the permutation sequence with the maximum length. This is because a longer permutation sequence indicates more uncertainty.

However, the probability assignment in Permutation distribution will not converge to a single element. Conducted from Definition 15, we have

$$\lim_{N \to \infty} P_{Per}(n|N) = \lim_{N \to \infty} \frac{P(N,n)}{\lfloor e \cdot N! \rfloor}$$

$$= \lim_{N \to \infty} \frac{N!}{(N-n)!} / (\lfloor e \cdot N! \rfloor)$$

$$= \lim_{N \to \infty} \frac{1}{e(N-n)!}.$$
(A13)

Thus, permutation distribution will converge to the limit form as shown in Table I. In contrast to the RPST distribution, the permutation distribution exhibits a greater degree of variability, which hinders the generation of i.i.d. random variables. Consequently, it is not suitable for simulating random walks.

3. Analysis on RPST-generated random walk

In previous section, it is proved that RPST distribution will converge to a probability distribution shown in Eq. (A12), ensuring its generation of i.i.d. random variables, which is a necessity for generating random walk. Some statistics properties of RPST-generated random walk are analyzed in this section.

The histogram in Fig. 2 displays the generation of random variables. It is expected that, for a fixed value of N, the RVG will produce random vectors that adhere to a normal distribution. This convergence towards a normal distribution is controlled by the central limit theorem (CLT) and Donsker's theorem, which ensure that as $N \to \infty$, the summation vector $\vec{V}_i = (V_x, V_y)$ will be distributed according to a normal distribution $N(0, \sigma_N^2)$, where $\sigma_N^2 = f(N)\sigma^2$ and $f(N) \propto N^2$ is a binomial function with respect to N. This result is supported by Lemma 2 and Fig. 3.

Due to the binomial growth of variance, we design the rescaling factor $\sqrt{\varrho}/(N\sqrt{N})$ to fitting the variance pattern of the Wiener process. This re-scaling factor comes from the following theorem:

Theorem 1. The RPST-generated random walk RW(t) can be converted to the Wiener process if the following limit form exists:

$$W(t) = \lim_{n,N\to\infty} RW_{n,N}(t) = \lim_{n,N\to\infty} \frac{\sqrt{\varrho}}{N\sqrt{N}} \frac{1}{\sqrt{n}} \sum_{1\le i\le \lfloor nt\rfloor} \vec{V}_i, t \in [0,1].$$
(A14)

Proof. The difference between RW(t) and Gaussian random walk lies in the variance, if we can scale the variance of RW(t) to fit into normal distribution, then following Definition 7 one can easily proves it.

As demonstrated in Lemmas 1 and 2, we get

$$\mathbb{E}\left[RW(t)\right] = \mathbb{E}\left[\sum_{i=0}^{t} V_i\right] = 0. \tag{A15}$$

Then we have

$$\mathbb{E}\left[RW^{2}(t)\right] = \sum_{i=1}^{t} \mathbb{E}\left[V_{i}^{2}\right] + 2\sum_{1 \le i < j < t} \mathbb{E}\left[V_{i}V_{j}\right]. \tag{A16}$$

Since random variables are independent with each other, then for any $i \neq j$, $\mathbb{E}\left[V_iV_j\right] = 0$.

Using the equation of variance

$$Var(X) = \mathbb{E}\left[X^2\right] - \mathbb{E}\left[X\right]^2,$$
 (A17)

we get

$$\mathbb{E}\left[V_i^2\right] = Var(V_i) \propto (N^2 + N). \tag{A18}$$

Together with Appendix 3, we get

$$\mathbb{E}\left[RW^{2}(t)\right] = \sum_{i=1}^{t} \mathbb{E}\left[V_{i}^{2}\right] = t \cdot (Var(V_{i})) \propto t(N^{2} + N). \quad (A19)$$

Finally, the variance of RW(t) has the following property:

$$Var(RW(t)) = \mathbb{E}\left[RW^2(t)\right] \propto t(N^2 + N).$$
 (A20)

Compared with Gaussian random walk, the growing speed of Var(RW(t)) is additionally multiplied by N^2+N . In other words, after t-s steps, the increments of Gaussian random walk $Z_{t-s} \sim N(0,(t-s)\sigma^2)$, while in random walk from RPST we have $Z_{t-s} \sim N(0,(t-s)(N^2+N)\sigma^2)$. Thus, based on the propagation property of variance, we construct $V_i'=V_i/N\sqrt{N}$ to offset the term (N^2+N) . But there still exists a coefficient between the scaling V_i and the step of Gaussian random walk (or Wiener process in limiting form), as shown in Fig. 8, so we design a variance control factor ϱ , then the final scaling factor would be $\sqrt{\varrho}/(N\sqrt{N})$. We set $\varrho=24$ to approximating Gaussian random walk.

So after scaling the RPST-generated random walk to Gaussian random walk, one can use Donsker's theorem and Definition 7 to construct the form in Theorem 1 to convert a RPST-generated random walk to a Wiener process, thus Theorem 1 is proved. $\hfill \Box$

Similar to a Wiener process, the limit scale form of random walk from RPST $RW_{n,N}(0)$ is also characterized by the following properties:

- $RW_{n,N}(0) = 0$. This prosperity is achieved by setting the starting point to zero $V_0 = 0$.
- $RW_{n,N}(t)$ has independent increments. This property is determined by the fact that each random variables are independent with each other, following a step size distribution of normal distribution $N(0, f(N)\sigma^2)$.
- For any $0 \le s < t$, the increments $RW_{n,N}(t) RW_{n,N}(s) \sim N(0, t s)$. This can be done by setting the variance factor $\varrho = 24$, as shown in Fig. 8.
- $RW_{n,N}(t)$ is almost surely continuous in t, this is ensured by Donsker's theorem as $N, n \to \infty$.

REFERENCES

- Ansari-Rad, M., Abdi, Y., and Arzi, E., "Monte Carlo random walk simulation of electron transport in dye-sensitized nanocrystalline solar cells: Influence of morphology and trap distribution," J. Phys. Chem. C 116, 3212–3218 (2012).
- Chen, L. and Deng, Y., "Entropy of random permutation set," Commun. Stat. Theory Methods 53(11), 4127–4146 (2024).
- Chen, X. and Deng, Y., "Evidential software risk assessment model on ordered frame of discernment," Expert Syst. Appl. 250, 123786 (2024).
- Contreras-Reyes, J. E. and Kharazmi, O., "Belief Fisher-Shannon information plane: Properties, extensions, and applications to time series analysis," Chaos, Solitons Fractals 177, 114271 (2023).
- Cui, H., Zhou, L., Li, Y., and Kang, B., "Belief entropy-of-entropy and its application in the cardiac interbeat interval time series analysis," Chaos, Solitons Fractals 155, 111736 (2022).
- Deng, J., Deng, Y., and Bo Yang, J., "Random permutation set reasoning," IEEE Trans. Pattern Anal. Mach. Intell. 1–12 (published online 2024).
- Deng, Y., "Random permutation set," Int. J. Comput. Commun. Control 17, 4542 (2022).
- Dempster, A. P., "Upper and lower probabilities induced by a multivalued mapping," Ann. Math. Stat. 38, 325–339 (1967).
- Einstein, A., Investigations on the Theory of the Brownian Movement (Courier Corporation, 1956).
- He, H. and Xiao, F., "A novel quantum Dempster's rule of combination for pattern classification," Inf. Sci. 671, 120617 (2024).
- Huang, Y., Xiao, F., Cao, Z., and Lin, C.-T., "Fractal belief Rényi divergence with its applications in pattern classification," IEEE Trans. Knowl. Data Eng. 1–16 (2023).
- Huang, Y., Xiao, F., Cao, Z., and Lin, C.-T., "Higher order fractal belief Rényi divergence with its applications in pattern classification," IEEE Trans. Pattern Anal. Mach. Intell. 45, 14709–14726 (2023).
- Kessing, R. K., Yang, P. -Y., Manmana, S. R., and Cao, J., "Long-range non-equilibrium coherent tunneling induced by fractional vibronic resonances," J. Phys. Chem. Lett. 13, 6831–6838 (2022).
- Kharazmi, O. and Contreras-Reyes, J. E., "Deng-Fisher information measure and its extensions: Application to Conway's game of life," Chaos, Solitons Fractals 174, 113871 (2023).
- Kharazmi, O. and Contreras-Reyes, J. E., "Belief inaccuracy information measures and their extensions," Fluctuation Noise Lett. 23, 2450041 (2024).
- Kharazmi, O., Contreras-Reyes, J. E., and Balakrishnan, N., "Jensen–Fisher information and Jensen–Shannon entropy measures based on complementary discrete distributions with an application to Conway's game of life," Physica D 453, 133822 (2023).
- Kharazmi, O., Jamali, H., and Contreras-Reyes, J. E., "Fisher information and its extensions based on infinite mixture density functions," Physica A 624, 128959 (2023).
- Lawler, G. F. and Limic, V., Random Walk: A Modern Introduction (Cambridge University Press, 2010), Vol. 123.
- Li, S. and Xiao, F., "Normal distribution based on maximum Deng entropy," Chaos, Solitons Fractals 167, 113057 (2023).
- Ortiz-Vilchis, P., Lei, M., and Ramirez-Arellano, A., "Reformulation of Deng information dimension of complex networks based on a sigmoid asymptote," Chaos, Solitons Fractals 180, 114569 (2024).
- Qiang, C., Deng, Y., and Cheong, K. H., "Information fractal dimension of mass function," Fractals 30, 2250110 (2022).
- Shafer, G., A Mathematical Theory of Evidence (Princeton University Press, 1976), Vol. 42.
- Shannon, C. E., "A mathematical theory of communication," Bell Syst. Tech. J. 27, 379–423 (1948).
- Thompson, C. J., Kienle, D. F., and Schwartz, D. K., "Enhanced facilitated diffusion of membrane-associating proteins under symmetric confinement," J. Phys. Chem. Lett. 13, 2901–2907 (2022).
- Tojo, C. and Argyrakis, P., "Correlated random walk in continuous space," Phys. Rev. E 54, 58 (1996).
- Wang, Y., Cao, X., Weng, T., Yang, H., and Gu, C., "A convex principle of search time for a multi-biased random walk on complex networks," Chaos, Solitons Fractals 147, 110990 (2021).

- Wang, Y., Li, Z., and Deng, Y., "A new orthogonal sum in random permutation set," Fuzzy Sets Syst. **490**, 109034 (2024).
- Wang, Z., Jusup, M., Guo, H., Shi, L., Geček, S., Anand, M., Perc, M., Bauch, C. T., Kurths, J., and Boccaletti, S. et al., "Communicating sentiment and outlook reverses inaction against collective risks," Proc. Natl. Acad. Sci. 117, 17650–17655 (2020).
- Wang, Z., Jusup, M., Shi, L., Lee, J.-H., Iwasa, Y., and Boccaletti, S., "Exploiting a cognitive bias promotes cooperation in social dilemma experiments," Nat. Commun. 9, 2954 (2018).
- Wang, Z., Jusup, M., Wang, R.-W., Shi, L., Iwasa, Y., Moreno, Y., and Kurths, J., "Onymity promotes cooperation in social dilemma experiments," Sci. Adv. 3, e1601444 (2017).
- Wang, Z., Mu, C., Hu, S., Chu, C., and Li, X., "Modelling the dynamics of regret minimization in large agent populations: A master equation approach," in *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, (IJCAI-22), July 2022*, edited by L. de. Raedt (International Joint Conferences on Artificial Intelligence Organization, 2022), pp. 534–540.
- Xiao, F., "Generalized quantum evidence theory," Appl. Intell. 53, 14329–14344 (2023).
- Xiao, F., "Quantum X-entropy in generalized quantum evidence theory," Inf. Sci. **643**, 119177 (2023).

- Xiao, F., Cao, Z., and Lin, C.-T., "A complex weighted discounting multisource information fusion with its application in pattern classification," IEEE Trans. Knowl. Data Eng. 35(8), 7609–7623 (2023).
- Xiao, F. and Pedrycz, W., "Negation of the quantum mass function for multisource quantum information fusion with its application to pattern classification," IEEE Trans. Pattern Anal. Mach. Intell. 45, 2054–2070 (2023).
- Xiao, F., Wen, J., and Pedrycz, W., "Generalized divergence-based decision making method with an application to pattern classification," IEEE Trans. Knowl. Data Eng. 35, 6941–6956 (2023).
- Zhan, T., Zhou, J., Li, Z., and Deng, Y., "Generalized information entropy and generalized information dimension," Chaos, Solitons Fractals 184, 114976 (2024).
- Zhang, L. and Xiao, F., "Belief Rényi divergence of divergence and its application in time series classification," IEEE Trans. Knowl. Data Eng. 36, 3670–3681 (2024).
- Zhao, T., Li, Z., and Deng, Y., "Information fractal dimension of Random Permutation Set," Chaos, Solitons Fractals 174, 113883 (2023).
- Zhao, T., Li, Z., and Deng, Y., "Linearity in Deng entropy," Chaos, Solitons Fractals 178, 114388 (2024).
- Zhou, J., Li, Z., Cheong, K. H., and Deng, Y., "Limit of the maximum random permutation set entropy," arXiv:2403.06206 (2024).