

Article

Jeffreys Divergence and Generalized Fisher Information Measures on Fokker–Planck Space–Time Random Field

Jiaxing Zhang 

School of Mathematics, Tianjin University, Tianjin 300350, China; 2017233010@tju.edu.cn; Tel.: +86-1512-750-2081

Abstract: In this paper, we present the derivation of Jeffreys divergence, generalized Fisher divergence, and the corresponding De Bruijn identities for space–time random field. First, we establish the connection between Jeffreys divergence and generalized Fisher information of a single space–time random field with respect to time and space variables. Furthermore, we obtain the Jeffreys divergence between two space–time random fields obtained by different parameters under the same Fokker–Planck equations. Then, the identities between the partial derivatives of the Jeffreys divergence with respect to space–time variables and the generalized Fisher divergence are found, also known as the De Bruijn identities. Later, at the end of the paper, we present three examples of the Fokker–Planck equations on space–time random fields, identify their density functions, and derive the Jeffreys divergence, generalized Fisher information, generalized Fisher divergence, and their corresponding De Bruijn identities.

Keywords: space–time random field; Fokker–Planck equations; differential entropy; Jeffreys divergence; fisher information; De Bruijn identities



Citation: Zhang, J. Jeffreys Divergence and Generalized Fisher Information Measures on Fokker–Planck Space–Time Random Field. *Entropy* **2023**, *25*, 1445. <https://doi.org/10.3390/e25101445>

Academic Editor: Jean-Pierre Gazeau

Received: 24 August 2023

Revised: 1 October 2023

Accepted: 11 October 2023

Published: 13 October 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

Information entropy and Fisher information are quantities to measure random information, and entropy divergence is derived from information entropy to measure the difference between two probability distributions. Formally, we can construct straightforward definitions of entropy divergence and Fisher information for the case of a space–time random field found on classical definitions. The density function, in their definitions, can be obtained in many different ways. In this paper, the density function of a space–time random field is obtained by Fokker–Planck equations. The traditional Fokker–Planck equation is a partial differential equation that describes the probability density function of a random process [1]. It describes the density function’s time-varying change rule. However, the Fokker–Planck equations for random fields, especially for space–time random fields, do not yet possess a distinct form. The classical equation needs to be generalized because the variable varies from time to space–time.

In this paper, we mainly obtain the relation between Jeffreys divergence and generalized Fisher information measure for space–time random field generated by Fokker–Planck equations. Jeffreys divergence is a symmetric entropy divergence, which is generalized from Kullback–Leibler divergence (KL divergence). Jeffreys divergence is a measure in information theory and statistics that evaluates the variation between anticipated and real probability distributions. However, if there is no overlap between the two distributions, the outcome will be infinite, which is a limitation of this approach. To prevent infinite results, we examine how Jeffreys divergence relates to generalized Fisher information for a space–time random field with slight variations in space–time parameters.

Moreover, the classical De Bruijn identity describes the relationship between differential entropy and the Fisher information of the Gaussian channel [2], and it can be generalized to other cases [3–7]. With gratitude to their works and following their ideas, we obtain De Bruijn identities on Jeffreys divergence and generalized Fisher information of space–time random fields, whose density functions satisfy Fokker–Planck equations.

1.1. Space–Time Random Field

The random field was first studied by Kolmogorov [8–10], and it was gradually improved by Yaglom [11–13] in the middle of the last century. A random field with $n \in \mathbb{N}_+$ variables can be expressed as

$$X(t_1, t_2, \dots, t_n) \quad (1)$$

where $(t_1, t_2, \dots, t_n) \in \mathbb{R}^n$. We call (1) a generalized random field or a multiparameter stochastic process. In some practical applications, we often use the concept of space–time random field. The space–time random field on a d -dimensional space is expressed as

$$X(t, x) \quad (2)$$

where $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^d$ are the space–time variables. It has many applications in statistics, finance, signal processing, stochastic partial differential equations, and other fields [14–27].

1.2. Kramers–Moyal Expansion and Fokker–Planck Equation

In the literature of stochastic processes, Kramers–Moyal expansion refers to a Taylor series of the master equation, named after Kramers and Moyal [28,29]. The Kramers–Moyal expansion is an infinite order partial differential equation

$$\frac{\partial}{\partial t} p(u, t) = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial u^n} [K_n(u, t) p(u, t)] \quad (3)$$

where $p(u, t)$ is the density function and

$$K_n(u, t) = \int_{\mathbb{R}} (u' - u)^n W(u'|u, t) du' \quad (4)$$

is the n -order conditional moment. Here, $W(u'|u, t)$ is the transition probability rate. The Fokker–Planck equation is obtained by keeping only the first two terms of the Kramers–Moyal expansion. In statistical mechanics, the Fokker–Planck equation is usually used to describe the time evolution of the probability density function of the velocity of a particle under the influence of drag forces and random forces, as in the famous Brownian motion, and this equation is commonly employed for determining the density function of an Itô stochastic differential equation [1].

1.3. Differential Entropy and De Bruijn Identity

The entropy of a continuous distribution was proposed by Shannon in 1948, known as differential entropy [30]:

$$h(X) = - \int_{\mathbb{R}^d} p(x) \log p(x) dx \quad (5)$$

where $h(\cdot)$ represents the differential entropy and $p(\cdot)$ is the probability density function of X . However, differential entropy is not easy to calculate and seldom exists. There are related studies on the entropy of stochastic processes and continuous systems [31–34]. If we consider a classical one-dimensional Gaussian channel model

$$Y_t = X + \sqrt{t}G \quad (6)$$

where X is the input signal, G is standard Gaussian noise, $t \geq 0$ is the strength, and Y_t is the output, we can obtain that the density of Y_t satisfies the following Fokker–Planck equation:

$$\frac{\partial}{\partial t} p(y, t) = \frac{1}{2} \frac{\partial^2}{\partial y^2} p(y, t) \quad (7)$$

Furthermore, the differential entropy of Y_t can be calculated, and then its derivative with respect to t can be obtained as

$$\frac{dh_{Y_t}(t)}{dt} = \frac{1}{2}FI_{Y_t}(t) \tag{8}$$

where

$$FI_{Y_t}(t) = \int_{\mathbb{R}} \left[\frac{\partial \log p(y, t)}{\partial y} \right]^2 p(y, t) dy \tag{9}$$

is the Fisher information of Y_t . The Equation (8) here is the De Burijn identity. The de Bruijn identity connects the differential entropy $h(\cdot)$ and the Fisher information $FI(\cdot)$, which shows that they are different aspects of the concept of “information”.

1.4. Entropy Divergence

In information theory and statistics, an entropy divergence is a statistical distance generated from information entropy to measure the difference between two probability distributions. There are various divergences generated by information entropy, such as Kullback–Leibler divergence [35], Jeffreys divergence [36], Jensen-Shannon divergence [37], and Rényi divergence [38]. These measures are applied in a variety of fields such as finance, economics, biology, signal processing, pattern recognition, and machine learning [39–49]. In this paper, we mainly focus on the Jeffreys divergence of two distributions, formed as

$$JD(P, Q) = \int_{\mathbb{R}} [p(u) - q(u)] \log \frac{p(u)}{q(u)} d\mu(u) \tag{10}$$

where μ is a measure of u .

2. Notations, Definitions, and Propositions

2.1. Notations and Assumptions

In this paper, we use the subsequent notations and definitions

- Given a probability space (Ω, \mathcal{F}, P) , two real valued space–time random fields are denoted as $X(\omega; t, x)$, $Y(\omega; s, y)$ or $X(t, x)$, $Y(s, y)$, where $\omega \in \Omega$ and (t, x) $(s, y) \in \mathbb{R}_+ \times \mathbb{R}^d$, $d \in \mathbb{N}_+$ are space–time variables.
- The probability density functions of P and Q are denoted as p and q . With $\forall u \in \mathbb{R}$, $p(u; t, x)$ is the density value at (t, x) of X and $q(u; s, y)$ is the density value at (s, y) of Y .
- Unless there are specific restrictions on the ranges of variables, suppose that our density functions $p(u; t, x)$ and $q(u, s, y)$ belongs to $C^{2,1,1}(\mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}^d, \mathbb{R})$. This means that $p(u; t, x)$ and $q(u; s, y)$ are partial differentiable twice with respect to u and once with respect to (t, x) or (s, y) , respectively.
- Vectors that differ only from the k -th coordinate of $x = (x_1, x_2, \dots, x_k, \dots, x_d)$ are denoted $\tilde{x}^{(k)} = (x_1, x_2, \dots, x'_k, \dots, x_d)$, where the k -th coordinates are x_k and x'_k , $k = 1, 2, \dots, d$.

2.2. Definitions

To obtain the generalized De Bruijn identities between Jeffreys divergence and Fisher divergence, we need to introduce some new definitions and propositions.

The primary and most important measure of information is the Kullback–Leibler divergence for random fields. Definition 1 is easily obtained as follows.

Definition 1. The Kullback–Leibler divergence between two space–time random fields $X(t, x)$ and $Y(s, y)$, (t, x) , $(s, y) \in \mathbb{R}_+ \times \mathbb{R}^d$, with density functions $p(u; t, x)$ and $q(u; s, y)$, is defined as

$$KL(P(t, x) \| Q(s, y)) = \int_{\mathbb{R}} p(u; t, x) \log \frac{p(u; t, x)}{q(u; s, y)} du \tag{11}$$

Similar to the classical Kullback–Leibler divergence, Kullback–Leibler divergence on random fields is not symmetrical, i.e.,

$$KL(P(t, x)||Q(s, y)) \neq KL(Q(s, y)||P(t, x)) \tag{12}$$

Following the classical definition of Jeffreys divergence on two random variables, we mainly consider Jeffreys divergence for random fields in this paper.

Definition 2. The Jeffreys divergence between space–time random fields $X(t, x)$ and $Y(s, y)$, $(t, x), (s, y) \in \mathbb{R}_+ \times \mathbb{R}^d$, with density function $p(u; t, x)$ and $q(u; s, y)$ is defined as

$$JD(P(t, x), Q(s, y)) = KL(P(t, x)||Q(s, y)) + KL(Q(s, y)||P(t, x)) \tag{13}$$

Here, we replace $\|$ with $,$ in the distortion measure to emphasize the symmetric property.

Another significant measure of information is Fisher information. In this paper, we consider the generalized Fisher information of the space–time random field.

Definition 3. The Generalized Fisher information of the space–time random field $X(t, x)$, $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^d$, with density function $p(u; t, x)$ defined by nonnegative function $f(\cdot)$, is formed as

$$FI_f(P(t, x)) = \int_{\mathbb{R}} f(u)[\partial_u \log p(u; t, x)]^2 p(u; t, x) du \tag{14}$$

In this case, where f is equal to 1, $FI_1(P(t, x))$ represents the typical Fisher information. In addition to Equation (14), there are similar forms of generalized Fisher information

$$FI_f^{(t)}(P(t, x)) = \int_{\mathbb{R}} f(u)[\partial_t \log p(u; t, x)]^2 p(u; t, x) du \tag{15}$$

and

$$FI_f^{(x_k)}(P(t, x)) = \int_{\mathbb{R}} f(u)[\partial_{x_k} \log p(u; t, x)]^2 p(u; t, x) du \tag{16}$$

for $k = 1, 2, \dots, d$.

Obviously, (15) and (16) are generalized Fisher information on space–time variables. Regarding the generalized Fisher information (14), we can come to a following simple proposition.

Proposition 1. For the arbitrary positive continuous function $f(\cdot)$, suppose the generalized Fisher information of continuous random variable X

$$FI_f(X) := \int_{\mathbb{R}} f(u) \left[\frac{d \log p_X(u)}{du} \right]^2 p_X(u) du \tag{17}$$

is well defined, where $p_X(u)$ represents the probability density. Then, we have the generalized Fisher information inequality

$$\frac{1}{FI_f(X + Y)} \geq \frac{1}{FI_f(X)} + \frac{1}{FI_f(Y)} \tag{18}$$

when $f \equiv 1$, $FI_1(X)$ represents the Fisher information in the standard case.

Proof. Denote $Z = X + Y$, p_X , p_Y , and p_Z represent densities, i.e.,

$$p_Z(z) = \int_{\mathbb{R}} p_X(x)p_Y(z - x)dx \tag{19}$$

and derivative function

$$p'_Z(z) = \int_{\mathbb{R}} p'_X(x)p_Y(z - x)dx \tag{20}$$

If $p_X, p_Y,$ and p_Z never vanish,

$$\begin{aligned} & \frac{p'_Z(z)}{p_Z(z)} \\ &= \int_{\mathbb{R}} \frac{p'_X(x)p_Y(z-x)}{p_Z(z)} dx \\ &= \int_{\mathbb{R}} \frac{p_X(x)p_Y(z-x)}{p_Z(z)} \frac{p'_X(x)}{p_X(x)} dx \\ &= E \left[\frac{p'_X(x)}{p_X(x)} \middle| Z = z \right] \end{aligned} \tag{21}$$

is the conditional expectation of $\frac{p'_X(x)}{p_X(x)}$ for given z . Similarly, we can obtain

$$\frac{p'_Z(z)}{p_Z(z)} = E \left[\frac{p'_Y(y)}{p_Y(y)} \middle| Z = z \right] \tag{22}$$

and $\forall \mu, \lambda \in \mathbb{R}$, we also find that

$$E \left[\mu \frac{p'_X(x)}{p_X(x)} + \lambda \frac{p'_Y(y)}{p_Y(y)} \middle| Z = z \right] = (\mu + \lambda) \frac{p'_Z(z)}{p_Z(z)} \tag{23}$$

Then, we have

$$\begin{aligned} & \left[(\mu + \lambda) \frac{p'_Z(z)}{p_Z(z)} \right]^2 \\ &= \left\{ E \left[\mu \frac{p'_X(x)}{p_X(x)} + \lambda \frac{p'_Y(y)}{p_Y(y)} \middle| Z = z \right] \right\}^2 \\ &\leq E \left\{ \left[\mu \frac{p'_X(x)}{p_X(x)} + \lambda \frac{p'_Y(y)}{p_Y(y)} \right]^2 \middle| Z = z \right\} \end{aligned} \tag{24}$$

with equality only if

$$\mu \frac{p'_X(x)}{p_X(x)} + \lambda \frac{p'_Y(y)}{p_Y(y)} = (\mu + \lambda) \frac{p'_Z(z)}{p_Z(z)} \tag{25}$$

with probability 1 whenever $z = x + y$ and we have

$$f(z) \left[(\mu + \lambda) \frac{p'_Z(z)}{p_Z(z)} \right]^2 \leq f(z) E \left[\left(\mu \frac{p'_X(x)}{p_X(x)} + \lambda \frac{p'_Y(y)}{p_Y(y)} \right)^2 \middle| Z = z \right] \tag{26}$$

Averaging both sides over the distribution of z

$$\begin{aligned} & E \left\{ f(z) \left[(\mu + \lambda) \frac{p'_Z(z)}{p_Z(z)} \right]^2 \right\} \\ &\leq E \left\{ f(z) E \left[\left(\mu \frac{p'_X(x)}{p_X(x)} + \lambda \frac{p'_Y(y)}{p_Y(y)} \right)^2 \middle| Z = z \right] \right\} \\ &= \mu^2 E \left\{ f(z) E \left[\left(\frac{p'_X(x)}{p_X(x)} \right)^2 \middle| Z = z \right] \right\} + \lambda^2 E \left\{ E \left[f(z) \left(\frac{p'_Y(y)}{p_Y(y)} \right)^2 \middle| Z = z \right] \right\} \end{aligned} \tag{27}$$

i.e.,

$$(\mu + \lambda)^2 FI_f(X + Y) \leq \mu^2 FI_f(X) + \lambda^2 FI_f(Y) \tag{28}$$

Let $\mu = \frac{1}{FI_f(X)}$ and $\lambda = \frac{1}{FI_f(Y)}$, we obtain

$$\frac{1}{FI_f(X+Y)} \geq \frac{1}{FI_f(X)} + \frac{1}{FI_f(Y)} \tag{29}$$

□

According to Definition 3, we can obtain relevant definitions on the generalized Fisher information measure.

Definition 4. The generalized Cross-Fisher information for space-time random fields $X(t, x)$ and $Y(s, y)$, $(t, x), (s, y) \in \mathbb{R}_+ \times \mathbb{R}^d$, with density functions $p(u; t, x)$ and $q(u; s, y)$, defined by the nonnegative function $f(\cdot)$, is defined as

$$CFI_f(P(t, x), Q(s, y)) = \int_{\mathbb{R}} f(u) [\partial_u \log q(u; s, y)]^2 p(u; t, x) du \tag{30}$$

Similar to the concept of cross-entropy, it is easy to verify that (30) is symmetrical about P and Q .

Definition 5. The generalized Fisher divergence for space-time random fields $X(t, x)$ and $Y(s, y)$, for $(t, x), (s, y) \in \mathbb{R}_+ \times \mathbb{R}^d$, with density functions $p(u; t, x)$ and $q(u; s, y)$, defined by nonnegative function $f(\cdot)$, is defined as

$$FD_f(P(t, x) \| Q(s, y)) = \int_{\mathbb{R}} f(u) [\partial_u \log p(u; t, x) - \partial_u \log q(u; s, y)]^2 p(u; t, x) du \tag{31}$$

In particular, when $f \equiv 1$, $FD_1(P(t, x) \| Q(s, y))$ represents the typical Fisher divergence.

Obviously, the generalized Fisher divergence between two random fields is not a symmetrical measure of information. We need to create a new formula to expand on (31) in order to achieve symmetry.

Definition 6. The generalized Fisher divergence for space-time random fields $X(t, x)$ and $Y(s, y)$, $(t, x), (s, y) \in \mathbb{R}_+ \times \mathbb{R}^d$, with density functions $p(u; t, x)$ and $q(u; s, y)$, defined by nonnegative functions $f(\cdot)$ and $g(\cdot)$, is defined as

$$FD_{(f,g)}(P(t, x) \| Q(s, y)) = \int_{\mathbb{R}} [f(u; t, x) \partial_u \log p(u; t, x) - g(u; s, y) \partial_u \log q(u; s, y)] \times [\partial_u \log p(u; t, x) - \partial_u \log q(u; s, y)] [p(u; t, x) + q(u; s, y)] du \tag{32}$$

In particular, if f equals g , the generalized Fisher divergence for random fields using a single function is denoted as $FD_{(f,f)}(P(t, x) \| Q(s, y))$.

In general, $FD_{(f,g)}(P(t, x) \| Q(s, y))$ is asymmetric with respect to P and Q , i.e.,

$$FD_{(f,g)}(P(t, x) \| Q(s, y)) \neq FD_{(f,g)}(Q(s, y) \| P(t, x)) \tag{33}$$

If we suppose that f and g are functions only related to P and Q , i.e.,

$$\begin{cases} f(u; t, x) = \mathcal{T}[p(t, x)](u) \\ g(u; s, y) = \mathcal{T}[q(s, y)](u) \end{cases} \tag{34}$$

where \mathcal{T} is an operator; the generalized Fisher divergence $FD_{(f,g)}(P(t, x) \| Q(s, y))$ can be rewritten as

$$\begin{aligned}
 & FD_{(f,g)}(P(t, x) \| Q(s, y)) \\
 &= \int_{\mathbb{R}} \{ \mathcal{T}[p(t, x)](u) \partial_u \log p(u; t, x) - \mathcal{T}[q(s, y)](u) \partial_u \log q(u; s, y) \} \\
 & \quad \times [\partial_u \log p(u; t, x) - \partial_u \log q(u; s, y)] [p(u; t, x) + q(u, s, y)] du
 \end{aligned} \tag{35}$$

and we can easily obtain

$$FD_{(f,g)}(P(t, x) \| Q(s, y)) = FD_{(g,f)}(Q(t, x) \| P(s, y)) \tag{36}$$

In this case, we call (35) symmetric Fisher divergence for random fields generated by operator \mathcal{T} and denote it as

$$\begin{aligned}
 & sFD_{\mathcal{T}}(P(t, x), Q(s, y)) \\
 &= \int_{\mathbb{R}} \{ \mathcal{T}[p(t, x)](u) \partial_u \log p(u; t, x) - \mathcal{T}[q(s, y)](u) \partial_u \log q(u; s, y) \} \\
 & \quad \times [\partial_u \log p(u; t, x) - \partial_u \log q(u; s, y)] [p(u; t, x) + q(u, s, y)] du
 \end{aligned} \tag{37}$$

Notice that

$$\begin{aligned}
 & Aa - Bb \\
 &= \frac{1}{2} \times 2(Aa - Bb) \\
 &= \frac{1}{2} [(Aa - Ab + Ab - Bb) + (Aa - Ba + Ba - Bb)] \\
 &= \frac{1}{2} [(A + B)(a - b) + (A - B)(a + b)]
 \end{aligned} \tag{38}$$

for $A, B, a, b \in \mathbb{R}$; then, we can rewrite (37) as

$$\begin{aligned}
 & sFD_{\mathcal{T}}(P(t, x), Q(s, y)) \\
 &= \frac{1}{2} \int_{\mathbb{R}} \{ \mathcal{T}[p(t, x)](u) + \mathcal{T}[q(s, y)](u) \} [\partial_u \log p(u; t, x) - \partial_u \log q(u; s, y)]^2 \\
 & \quad \times [p(u; t, x) + q(u, s, y)] du \\
 & \quad + \frac{1}{2} \int_{\mathbb{R}} \{ \mathcal{T}[p(t, x)](u) - \mathcal{T}[q(s, y)](u) \} \{ [\partial_u \log p(u; t, x)]^2 - [\partial_u \log q(u; s, y)]^2 \} \\
 & \quad \times [p(u; t, x) + q(u, s, y)] du \\
 &= \frac{1}{2} [FD_{\mathcal{T}[p(t,x)]+\mathcal{T}[q(s,y)]}(P(t, x) \| Q(s, y)) + FD_{\mathcal{T}[p(t,x)]+\mathcal{T}[q(s,y)]}(Q(s, y) \| P(t, x))] \\
 & \quad + \frac{1}{2} [FI_{\mathcal{T}[p(t,x)]-\mathcal{T}[q(s,y)]}(P(t, x)) + FI_{\mathcal{T}[p(t,x)]-\mathcal{T}[q(s,y)]}(Q(s, y))] \\
 & \quad + \frac{1}{2} [CFI_{\mathcal{T}[p(t,x)]-\mathcal{T}[q(s,y)]}(Q(s, y), P(t, x)) - CFI_{\mathcal{T}[p(t,x)]-\mathcal{T}[q(s,y)]}(P(t, x), Q(s, y))]
 \end{aligned} \tag{39}$$

Lemma 1 (Kramers–Moyal expansion [28,29]). Suppose that the random process $X(t)$ has any order moment; then, the probability density function $p(u, t)$ satisfies the Kramers–Moyal expansion

$$\frac{\partial}{\partial t} p(u, t) = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial u^n} [K_n(u, t) p(u, t)] \tag{40}$$

where

$$K_n(u, t) = \int_{\mathbb{R}} (u' - u)^n W(u' | u, t) du' \tag{41}$$

is the n -order conditional moment and $W(u' | u, t)$ is the transition probability rate.

Lemma 2 (Pawula theorem [50,51]). *If the limit on conditional moment of random process $X(t)$*

$$\lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} E \left\{ [X(t + \Delta t) - X(t)]^n \mid X(t) = x \right\} \tag{42}$$

exists for all $n \in \mathbb{N}_+$, and the limit value equals 0 for some even number, then the limit values are 0 for all $n \geq 3$.

The Pawula theorem states that there are only three possible cases in the Kramers–Moyal expansion:

- (1) The Kramers–Moyal expansion is truncated at $n = 1$, meaning that the process is deterministic;
- (2) The Kramers–Moyal expansion stops at $n = 2$, with the resulting equation being the Fokker–Planck equation, and describes diffusion processes;
- (3) The Kramers–Moyal expansion contains all the terms up to $n = \infty$.

In this paper, we only focus on the case of the Fokker–Planck equation.

3. Main Results and Proofs

In this section, we establish the Fokker–Planck equations for continuous space–time random field. Additionally, we present the relationship theorem between Jeffreys divergence and Fisher information, as well as the De Bruijn identities connection between Jeffreys divergence and Fisher divergence.

Theorem 1. *The probability density function $p(u; t, x)$ of the continuous space–time random field $X(t, x)$, $u \in \mathbb{R}$, $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^d$ satisfies the following Fokker–Planck equations:*

$$\begin{cases} \frac{\partial}{\partial t} p(u; t, x) = \frac{1}{2} \frac{\partial^2}{\partial u^2} [b_0(u; t, x) p(u, t, x)] - \frac{\partial}{\partial u} [a_0(u; t, x) p(u; t, x)] \\ \frac{\partial}{\partial x_k} p(u; t, x) = \frac{1}{2} \frac{\partial^2}{\partial u^2} [b_k(u; t, x) p(u, t, x)] - \frac{\partial}{\partial u} [a_k(u; t, x) p(u; t, x)] \\ k = 1, 2, \dots, d \end{cases} \tag{43}$$

where

$$\begin{cases} a_0(u; t, x) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} M_1(u; t, \Delta t, x) \\ b_0(u; t, x) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} M_2(u; t, \Delta t, x) \\ a_k(u; t, x) = \lim_{\Delta x_k \rightarrow 0} \frac{1}{\Delta x_k} \tilde{M}_1(u; t, x, \Delta x_k) \\ b_k(u; t, x) = \lim_{\Delta x_k \rightarrow 0} \frac{1}{\Delta x_k} \tilde{M}_2(u; t, x, \Delta x_k) \\ k = 1, 2, \dots, d \end{cases} \tag{44}$$

here,

$$\begin{cases} M_n(u; t, \Delta t, x) = E \left\{ [X(t + \Delta t, x) - X(t, x)]^n \mid X(t, x) = u \right\} \\ \tilde{M}_n(u; t, x, \Delta x_k) = E \left\{ [X(t, x + \Delta x_k e_k) - X(t, x)]^n \mid X(t, x) = u \right\} \end{cases} \tag{45}$$

are n -order conditional moments and $e_k = (0, 0, \dots, 1, \dots, 0) \in \mathbb{R}^d$ are standard orthogonal basis vectors, $k = 1, 2, \dots, d$.

Proof. $\forall \Delta t \neq 0$, we can obtain the difference of density function in the time variable

$$p(u; t + \Delta t, x) - p(u; t, x) = \sum_{n=1}^{+\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial u^n} [M_n(u; t, \Delta t, x) p(u; t, x)] \tag{46}$$

where

$$M_n(u; t, \Delta t, x) = E \left\{ [X(t + \Delta t, x) - X(t, x)]^n \mid X(t, x) = u \right\} \tag{47}$$

is the n -order conditional moment. Then, the partial derivative of the density function with respect to t is

$$\frac{\partial}{\partial t} p(u; t, x) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \sum_{n=1}^{+\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial u^n} [M_n(u; t, \Delta t, x) p(u; t, x)] \tag{48}$$

The Pawula theorem implies that if the Kramers–Moyal expansion stops after the second term, we obtain the Fokker–Planck equation about the time variable t

$$\frac{\partial}{\partial t} p(u; t, x) = \frac{1}{2} \frac{\partial^2}{\partial u^2} [b_0(u; t, x) p(u, t, x)] - \frac{\partial}{\partial u} [a_0(u; t, x) p(u; t, x)] \tag{49}$$

where

$$\begin{cases} a_0(u; t, x) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} M_1(u; t, \Delta t, x) \\ b_0(u; t, x) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} M_2(u; t, \Delta t, x) \end{cases} \tag{50}$$

Similarly, we may consider the increment Δx_k of the spatial variable x_k , and we can obtain the Fokker–Planck equations about x_k as

$$\frac{\partial}{\partial x_k} p(u; t, x) = \frac{1}{2} \frac{\partial^2}{\partial u^2} [b_k(u; t, x) p(u, t, x)] - \frac{\partial}{\partial u} [a_k(u; t, x) p(u; t, x)] \tag{51}$$

where

$$\begin{cases} a_k(u; t, x) = \lim_{\Delta x_k \rightarrow 0} \frac{1}{\Delta x_k} \tilde{M}_1(u; t, x, \Delta x_k) \\ b_k(u; t, x) = \lim_{\Delta x_k \rightarrow 0} \frac{1}{\Delta x_k} \tilde{M}_2(u; t, x, \Delta x_k) \end{cases} \tag{52}$$

here,

$$\tilde{M}_n(u; t, x, \Delta x_k) = E \left\{ [X(t, x + \Delta x_k e_k) - X(t, x)]^n \mid X(t, x) = u \right\} \tag{53}$$

$e_k = (0, 0, \dots, 1, \dots, 0) \in \mathbb{R}^d$ are standard orthogonal basis vectors, $k = 1, 2, \dots, d$. □

The Fokker–Planck equations are partial differential equations that describe the probability density function of the space–time random field, similar to the classical Fokker–Planck equation. Solving a system of partial differential equations for general Fokker–Planck equations proves to be challenging. Fortunately, in Section 4 we present three distinct categories of space–time random fields in detail, along with their corresponding Fokker–Planck equations, and deduce their probability density functions.

Next, we examine the relationship between Jeffreys divergence and Fisher information in a single space–time random field when there are different time or spatial variables.

Theorem 2. Suppose that $p(u; t, x) > 0$ is a continuous differential density function of the space–time random field $X(t, x)$, the partial derivatives $\partial_u p(u; t, x)$, $\partial_t p(u; t, x)$, $\partial_{x_k} p(u; t, x)$ are continuous bounded functions, and the integrals in the proof are well-defined, $k = 1, 2, \dots, d$, $u \in \mathbb{R}$, $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^d$. Then, we have

$$\begin{cases} \lim_{|t-s| \rightarrow 0} \frac{JD(P(t, x), P(s, x))}{|t-s|^2} = FI_1^{(t)}(X(t, x)) \\ \lim_{|x_k-x'_k| \rightarrow 0} \frac{JD(P(t, x), P(t, \tilde{x}^{(k)}))}{|x_k-x'_k|^2} = FI_1^{(x_k)}(X(t, x)) \end{cases} \tag{54}$$

$k = 1, 2, \dots, d$

Proof. For fixed $x \in \mathbb{R}$, and $\forall s \neq t > 0$,

$$\begin{aligned} JD(P(t, x), P(s, x)) &= KL(P(t, x) || P(s, x)) + KL(P(s, x) || P(t, x)) \\ &= \int_{\mathbb{R}} [\log p(u; t, x) - \log p(u, s, x)] [p(u; t, x) - p(u, s, x)] du \end{aligned} \tag{55}$$

then we can obtain

$$\begin{aligned} &\lim_{|t-s| \rightarrow 0} \frac{JD(P(t, x), P(s, x))}{|t-s|^2} \\ &= \lim_{|t-s| \rightarrow 0} \int_{\mathbb{R}} \frac{\log p(u; t, x) - \log p(u, s, x)}{t-s} \frac{p(u; t, x) - p(u, s, x)}{t-s} du \end{aligned} \tag{56}$$

Notice that

$$\begin{cases} \lim_{|t-s| \rightarrow 0} \frac{\log p(u; t, x) - \log p(u, s, x)}{t-s} = \partial_t \log p(u; t, x) \\ \lim_{|t-s| \rightarrow 0} \frac{p(u; t, x) - p(u, s, x)}{t-s} = \partial_t p(u; t, x) \end{cases} \tag{57}$$

exist, and we obtain

$$\begin{aligned} &\lim_{|t-s| \rightarrow 0} \frac{JD(P(t, x), P(s, x))}{|t-s|^2} \\ &= \int_{\mathbb{R}} \partial_t \log p(u; t, x) \partial_t p(u; t, x) du \\ &= \int_{\mathbb{R}} \partial_t [\log p(u; t, x)]^2 p(u; t, x) du \\ &= FI_1^{(t)}(X(t, x)) \end{aligned} \tag{58}$$

Similarly, for fixed t and $\forall x_k \neq x'_k$, we can obtain the identity on Jeffreys divergence and Fisher information for space coordinates

$$\lim_{|x_k - x'_k| \rightarrow 0} \frac{JD(P(t, x), P(t, \tilde{x}^{(k)}))}{|x_k - x'_k|^2} = FI_1^{(x_k)}(X(t, x)) \tag{59}$$

for $k = 1, 2, \dots, d$. \square

Theorem 2 states that as the space–time variable difference approaches zero, the Fisher information of the space–time random field is the limit of the ratio of Jeffreys divergence at different locations to the square of space–time variable difference. It is noteworthy that Theorem 2 specifically addresses Jeffreys divergence only in cases where a single space–time random field is situated in distinct space–time positions, and where the difference between space–time variables approaches to 0. This ensures that Jeffreys divergence will not be infinite.

Theorem 3. Suppose that $p(u; t, x)$ and $q(u; t, x)$ are continuous differentiable density functions of space–time random fields $X(t, x)$ and $Y(t, x)$ such that

$$\begin{cases} \lim_{u \rightarrow \infty} \left\{ \frac{1}{2} \partial_u [b_k^{(1)}(u; t, x) p(u; t, x)] - a_k^{(1)}(u; t, x) p(u; t, x) \right\} \left[\log \frac{p(u; t, x)}{q(u; t, x)} - \frac{q(u; t, x)}{p(u; t, x)} \right] = 0 \\ \lim_{u \rightarrow \infty} \left\{ \frac{1}{2} \partial_u [b_k^{(2)}(u; t, x) q(u; t, x)] - a_k^{(2)}(u; t, x) q(u; t, x) \right\} \left[\log \frac{q(u; t, x)}{p(u; t, x)} - \frac{p(u; t, x)}{q(u; t, x)} \right] = 0 \end{cases} \tag{60}$$

where a_k, b_k are the forms in (44) and (45), and $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^d, k = 0, 1, 2, \dots, d$. Then, the Jeffreys divergence $JD(P(t, x), Q(t, x))$ satisfies generalized De Bruijn identities

$$\begin{cases} \frac{\partial}{\partial t} JD(P(t, x), Q(t, x)) = -\frac{1}{2} FD_{(b_0^{(1)}, b_0^{(2)})} (P(t, x) \| Q(t, x)) - R_0(P(t, x) \| Q(t, x)) \\ \frac{\partial}{\partial x_k} JD(P(t, x), Q(t, x)) = -\frac{1}{2} FD_{(b_k^{(1)}, b_k^{(2)})} (P(t, x) \| Q(t, x)) - R_k(P(t, x) \| Q(t, x)) \\ k = 1, 2, \dots, d \end{cases} \tag{61}$$

where

$$\begin{cases} R_0(P(t, x) \| Q(t, x)) = \int_{\mathbb{R}} \left[\frac{1}{2} \partial_{uu}^2 (b_0^{(1)} - b_0^{(2)}) - \partial_u (a_0^{(1)} - a_0^{(2)}) \right] (p + q) du \\ R_k(P(t, x) \| Q(t, x)) = \int_{\mathbb{R}} \left[\frac{1}{2} \partial_{uu}^2 (b_k^{(1)} - b_k^{(2)}) - \partial_u (a_k^{(1)} - a_k^{(2)}) \right] (p + q) du \\ k = 1, 2, \dots, d \end{cases} \tag{62}$$

here, we omit $(u; t, x)$ in the integrals for convenience.

Proof. By Definition 2, we have

$$\begin{aligned} & JD(P(t, x), Q(t, x)) \\ &= KL(P(t, x) \| Q(t, x)) + KL(Q(t, x) \| P(t, x)) \\ &= \int_{\mathbb{R}} p \log \frac{p}{q} du + \int_{\mathbb{R}} q \log \frac{q}{p} du \\ &= \int_{\mathbb{R}} \left[p \log \frac{p}{q} + q \log \frac{q}{p} \right] du \end{aligned} \tag{63}$$

where $p := p(u; t, x)$, $q := q(u; t, x)$ are density functions of $X(t, x)$ and $Y(t, x)$; here, we omit $(u; t, x)$.

Notice that

$$\begin{cases} \partial_u \left(\frac{p}{q} \right) = \frac{1}{q} \left(\partial_u p - \frac{p}{q} \partial_u q \right) \\ \partial_u \left(\frac{q}{p} \right) = \frac{1}{p} \left(\partial_u q - \frac{q}{p} \partial_u p \right) \end{cases} \tag{64}$$

i.e.,

$$\begin{cases} \frac{p}{q} \partial_u q = \partial_u p - q \partial_u \left(\frac{p}{q} \right) \\ \frac{q}{p} \partial_u p = \partial_u q - p \partial_u \left(\frac{q}{p} \right) \end{cases} \tag{65}$$

and

$$\begin{aligned} & (\partial_u \log p - \partial_u \log q)(p + q) \\ &= \left(\frac{1}{p} \partial_u p - \frac{1}{q} \partial_u q \right) (p + q) \\ &= \partial_u p - \partial_u q + \frac{q}{p} \partial_u p - \frac{p}{q} \partial_u q \\ &= p \frac{q \partial_u p - p \partial_u q}{p^2} + q \frac{q \partial_u p - p \partial_u q}{q^2} \\ &= - \left[p \partial_u \left(\frac{q}{p} \right) - q \partial_u \left(\frac{p}{q} \right) \right] \end{aligned} \tag{66}$$

then,

$$\begin{aligned}
 & \partial_t JD(P(t, x), Q(t, x)) \\
 &= \int_{\mathbb{R}} \left[\partial_t p \log \frac{p}{q} + q \partial_t \left(\frac{p}{q} \right) + \partial_t q \log \frac{q}{p} + p \partial_t \left(\frac{q}{p} \right) \right] du \\
 &= \int_{\mathbb{R}} \left[\partial_t p \log \frac{p}{q} + \partial_t p - \frac{p}{q} \partial_t q + \partial_t q \log \frac{q}{p} + \partial_t q - \frac{q}{p} \partial_t p \right] du \\
 &= \int_{\mathbb{R}} \left[\left(\log \frac{p}{q} - \frac{q}{p} \right) \partial_t p + \left(\log \frac{q}{p} - \frac{p}{q} \right) \partial_t q \right] du \\
 &= \int_{\mathbb{R}} \left(\log \frac{p}{q} - \frac{q}{p} \right) \left[\frac{1}{2} \partial_{uu}^2 (b_0^{(1)} p) - \partial_u (a_0^{(1)} p) \right] du \\
 &\quad + \int_{\mathbb{R}} \left(\log \frac{q}{p} - \frac{p}{q} \right) \left[\frac{1}{2} \partial_{uu}^2 (b_0^{(2)} q) - \partial_u (a_0^{(2)} q) \right] du \\
 &= - \int_{\mathbb{R}} \left[\frac{1}{2} \partial_u (b_0^{(1)} p) - a_0^{(1)} p \right] \left[\frac{q}{p} \partial_u \left(\frac{p}{q} \right) - \partial_u \left(\frac{q}{p} \right) \right] du \\
 &\quad - \int_{\mathbb{R}} \left[\frac{1}{2} \partial_u (b_0^{(2)} q) - a_0^{(2)} q \right] \left[\frac{p}{q} \partial_u \left(\frac{q}{p} \right) - \partial_u \left(\frac{p}{q} \right) \right] du \\
 &= - \int_{\mathbb{R}} \left[\frac{1}{2} \partial_u (b_0^{(1)} p) - a_0^{(1)} p \right] \frac{1}{p} \left[q \partial_u \left(\frac{p}{q} \right) - p \partial_u \left(\frac{q}{p} \right) \right] du \\
 &\quad - \int_{\mathbb{R}} \left[\frac{1}{2} \partial_u (b_0^{(2)} q) - a_0^{(2)} q \right] \frac{1}{q} \left[p \partial_u \left(\frac{q}{p} \right) - q \partial_u \left(\frac{p}{q} \right) \right] du \\
 &= - \int_{\mathbb{R}} \left\{ \left[\frac{1}{2} \partial_u (b_0^{(1)} p) - a_0^{(1)} p \right] \frac{1}{p} - \left[\frac{1}{2} \partial_u (b_0^{(2)} q) - a_0^{(2)} q \right] \frac{1}{q} \right\} \\
 &\quad \times (\partial_u \log p - \partial_u \log q) (p + q) du \\
 &= - \int_{\mathbb{R}} \left[\frac{1}{2} b_0^{(1)} \partial_u \log p - \frac{1}{2} b_0^{(2)} \partial_u \log q + \frac{1}{2} \partial_u (b_0^{(1)} - b_0^{(2)}) - (a_0^{(1)} - a_0^{(2)}) \right] \\
 &\quad \times (\partial_u \log p - \partial_u \log q) (p + q) du \\
 &= - \int_{\mathbb{R}} \left[\frac{1}{2} b_0^{(1)} \partial_u \log p - \frac{1}{2} b_0^{(2)} \partial_u \log q \right] (\partial_u \log p - \partial_u \log q) (p + q) du \\
 &\quad - \int_{\mathbb{R}} \left[\frac{1}{2} \partial_u (b_0^{(1)} - b_0^{(2)}) - (a_0^{(1)} - a_0^{(2)}) \right] (\partial_u \log p - \partial_u \log q) (p + q) du \\
 &= - \frac{1}{2} FD_{(b_0^{(1)}, b_0^{(2)})} (P(t, x) \| Q(t, x)) - R_0(P(t, x) \| Q(t, x))
 \end{aligned} \tag{67}$$

where

$$\begin{aligned}
 & FD_{(b_0^{(1)}, b_0^{(2)})} (P(t, x) \| Q(t, x)) \\
 &= \int_{\mathbb{R}} \left[b_0^{(1)} \partial_u \log p - b_0^{(2)} \partial_u \log q \right] (\partial_u \log p - \partial_u \log q) (p + q) du
 \end{aligned} \tag{68}$$

and

$$R_0(P(t, x) \| Q(t, x)) = \int_{\mathbb{R}} \left[\frac{1}{2} \partial_{uu}^2 (b_0^{(1)} - b_0^{(2)}) - \partial_u (a_0^{(1)} - a_0^{(2)}) \right] (p + q) du \tag{69}$$

Similarly, for $k = 1, 2, \dots, d$, we can obtain the generalized De Bruijn identities about the spatial variable x_k

$$\frac{\partial}{\partial x_k} JD(P(t, x), Q(t, x)) = - \frac{1}{2} FD_{(b_k^{(1)}, b_k^{(2)})} (P(t, x) \| Q(t, x)) - R_k(P(t, x), Q(t, x)) \tag{70}$$

where

$$R_k(P(t, x) \| Q(t, x)) = \int_{\mathbb{R}} \left[\frac{1}{2} \partial_{uu}^2 (b_k^{(1)} - b_k^{(2)}) - \partial_u (a_k^{(1)} - a_k^{(2)}) \right] (p + q) du \tag{71}$$

then we obtain the conclusion. \square

Unlike Theorem 3, Theorem 4 focuses on the Jeffreys divergence between two separate space–time random fields $X(t, x)$ and $Y(t, x)$, both at the same position (t, x) , and establishes the identities of the connection between the Jeffreys divergence and the Fisher divergence of $X(t, x)$ and $Y(t, x)$. This is known as the De Bruijn identities. To prevent Jeffreys divergence from becoming infinite, it is necessary for the difference between the probability density functions of $X(t, x)$ and $Y(t, x)$ to be small. In Section 4, we obtain Jeffreys divergence and Fisher divergence using the same type of Fokker–Planck equations but with different parameters. This allows for the selection of only the appropriate parameters.

4. Three Fokker–Planck Random Fields and Their Corresponding Information Measures

In this section, we present three types of Fokker–Planck equations and derive their corresponding density functions and information measures, which are Jeffreys divergence, generalized Fisher information, and Fisher divergence. With these quantities, the results corresponding to the applications of Theorems 2 and 3 are obtained. On the one hand, we calculate the ratio of Jeffreys divergence to the square of space–time variation on the identical Fokker–Planck space–time random field at various space–time points, in comparison to generalized Fisher information. On the other hand, we derive the De Bruijn identities for Jeffreys divergence and generalized Fisher divergence from Fokker–Planck equations on a single space–time random field at the corresponding space–time location, under same type but with different parameters.

First, we present a theorem regarding simple type Fokker–Planck equations of the random field.

Theorem 4. *Suppose the functions in the Fokker–Planck Equations (43) for the continuous random field $X(t, x)$ are formulated as follows:*

$$\begin{cases} a_0(u; t, x) = a_0(t, x) \\ b_0(u; t, x) = b_0(t, x) > 0 \\ a_k(u; t, x) = a_k(t, x) \\ b_k(u; t, x) = b_k(t, x) > 0 \\ k = 1, 2, \dots, d \end{cases} \tag{72}$$

where $a_0, a_k, b_0,$ and b_k are continuously differentiable functions independent of u and two continuously differentiable functions, $\alpha(t, x)$ and $\beta(t, x)$, exist such that

$$\begin{cases} d\alpha(t, x) = a_0 dt + a_1 dx_1 + \dots + a_d dx_d \\ d\beta(t, x) = b_0 dt + b_1 dx_1 + \dots + b_d dx_d \end{cases} \tag{73}$$

the initial density function is $p(u; t, x) = \delta[u - u_0(x)]$ as $prod(t, x) = 0$; then, the density function of $X(t, x)$ is presented as follows:

$$p(u; t, x) = \frac{1}{\sqrt{2\pi\beta(t, x)}} e^{-\frac{[u - u_0(x) - \alpha(t, x)]^2}{2\beta(t, x)}} \tag{74}$$

Proof. It can be easily inferred that the Fokker–Planck equations are simple parabolic equations, and their solution can be obtained through Fourier transform

$$\begin{cases} p(u; t, x) = \frac{1}{\sqrt{2\pi \int_0^t b_0(s, x) ds}} e^{-\frac{[u-u_0(x)-\int_0^t a_0(s, x) ds]^2}{2 \int_0^t b_0(s, x) ds}} \\ p(u; t, x) = \frac{1}{\sqrt{2\pi \int_0^{x_k} b_k(t, x) dx_k}} e^{-\frac{[u-u_0(x)-\int_0^{x_k} a_k(t, x) dx_k]^2}{2 \int_0^{x_k} b_k(t, x) dx_k}} \end{cases} \tag{75}$$

Recall that there are two functions $\alpha(t, x)$ and $\beta(t, x)$ such that

$$\begin{cases} d\alpha(t, x) = a_0(t, x)dt + a_1(t, x)dx_1 + \dots + a_d(t, x)dx_d \\ d\beta(t, x) = b_0(t, x)dt + b_1(t, x)dx_1 + \dots + b_d(t, x)dx_d \end{cases} \tag{76}$$

we can obtain the probability density function

$$p(u; t, x) = \frac{1}{\sqrt{2\pi\beta(t, x)}} e^{-\frac{[u-u_0(x)-\alpha(t, x)]^2}{2\beta(t, x)}} \tag{77}$$

□

Actually, numerous examples exist in which the Fokker–Planck equations comply with Theorem 4. Let $B(t, x)$ be the $(1 + d, 1)$ Brownian sheet [52,53], that is, a centered continuous Gaussian process that is indexed by $(1 + d)$ real, positive parameters and takes its values in \mathbb{R} . Moreover, its covariance structure is given by

$$E[B(t, x)B(s, y)] = (t \wedge s) \prod_{k=1}^d (x_k \wedge y_k) \tag{78}$$

for $(t, x_1, x_2, \dots, x_d), (s, y_1, y_2, \dots, y_d) \in \mathbb{R}_+ \times \mathbb{R}_+^d$, where $(\cdot \wedge \cdot)$ represents the minimum of two numbers. We can easily obtain

$$E[B^2(t, x)] = \text{prod}(t, x) \tag{79}$$

where $\text{prod}(t, x) = tx_1x_2 \dots x_d$ is the coordinate product of (t, x) and the density function is

$$p^{(1)}(u; t, x) = \frac{1}{\sqrt{2\pi\text{prod}(t, x)}} e^{-\frac{u^2}{2\text{prod}(t, x)}} \tag{80}$$

Moreover, the Fokker–Planck equations of Brownian sheet are

$$\begin{cases} \frac{\partial}{\partial t} p^{(1)}(u; t, x) = \frac{\text{prod}(x)}{2} \frac{\partial^2}{\partial u^2} p^{(1)}(u, t, x) \\ \frac{\partial}{\partial x_k} p^{(1)}(u; t, x) = \frac{\text{prod}(t, x)}{2x_k} \frac{\partial^2}{\partial u^2} p^{(1)}(u, t, x) \\ k = 1, 2, \dots, d \end{cases} \tag{81}$$

with the initial condition $p(u; t, x) = \delta(u)$ as $\text{prod}(t, x) = 0$.

Following the concept of constructing a Brownian bridge on Brownian motion [53], we refer to

$$B^*(t, x) = B(t, x) - \text{prod}(t, x)B(1, 1, \dots, 1) \tag{82}$$

as a Brownian sheet bridge on the cube $(t, x) \in [0, 1] \times [0, 1]^d$, where $B(t, x)$ represents the Brownian sheet. Obviously, $B^*(t, x)$ is Gaussian, and $E[B^*(t, x)] = 0$ and the covariance structure are

$$E[B^*(t, x)B^*(s, y)] = E[B(t, x)B(s, y)] - \text{prod}(t, x)\text{prod}(s, y) \tag{83}$$

we can obtain

$$E \left[B^2(t, x) \right] = \text{prod}(t, x) [1 - \text{prod}(t, x)] \tag{84}$$

and the density function of $B^*(t, x)$ is

$$p^{(2)}(u; t, x) = \frac{1}{\sqrt{2\pi \text{prod}(t, x) [1 - \text{prod}(t, x)]}} e^{-\frac{u^2}{2\text{prod}(t, x)[1 - \text{prod}(t, x)]}} \tag{85}$$

In addition to this, the Fokker–Planck equations of Brownian sheet bridge are

$$\begin{cases} \frac{\partial}{\partial t} p^{(2)}(u; t, x) = \frac{\text{prod}(x)[1 - 2\text{prod}(t, x)]}{2} \frac{\partial^2}{\partial u^2} p^{(2)}(u, t, x) \\ \frac{\partial}{\partial x_k} p^{(2)}(u; t, x) = \frac{\text{prod}(t, x)}{2x_k} [1 - 2\text{prod}(t, x)] \frac{\partial^2}{\partial u^2} p^{(2)}(u, t, x) \\ k = 1, 2, \dots, d \end{cases} \tag{86}$$

with the initial condition $p(u; t, x) = \delta(u)$ as $\text{prod}(t, x) = 0$, and we obtain the solution (85).

Combining two probability density functions (80) and (85) yields their respective Jeffreys divergence and generalized De Burijn identities. The Jeffreys divergence of (74) can be obtained at various space–time points as

$$JD(P(t, x), P(s, y)) = \frac{[\alpha(t, x) - \alpha(s, y)]^2 + \beta(s, y)}{2\beta(t, x)} + \frac{[\alpha(t, x) - \alpha(s, y)]^2 + \beta(t, x)}{2\beta(s, y)} - 1 \tag{87}$$

and the Fisher divergence between $P^{(1)}$ and $P^{(2)}$ at the identical space–time point represents

$$\begin{aligned} & FD_{(b_k^{(1)}, b_k^{(2)})} \left(P^{(1)}(t, x) \| P^{(2)}(t, x) \right) \\ &= \frac{1}{\beta_1^2(t, x) \beta_2^2(t, x)} \left\{ [\alpha_1(t, x) - \alpha_2(t, x)]^2 [b_k^{(2)} \beta_1^2(t, x) + b_k^{(1)} \beta_2^2(t, x)] \right. \\ &\quad \left. + [\beta_1(t, x) - \beta_2(t, x)] [\beta_1(t, x) + \beta_2(t, x)] [b_k^{(2)} \beta_1(t, x) - b_k^{(1)} \beta_2(t, x)] \right\} \end{aligned} \tag{88}$$

where $k = 0, 1, \dots, d$.

Bring the density function of Brownian sheet into Equation (87); we can easily obtain the Jeffreys divergence of the Brownian sheet at various space–time points as

$$JD \left(P^{(1)}(t, x), P^{(1)}(s, y) \right) = \frac{\text{prod}(s, y)}{2\text{prod}(t, x)} + \frac{\text{prod}(t, x)}{2\text{prod}(s, y)} - 1 \tag{89}$$

and the generalized Fisher information on space–time variables is as follows:

$$\begin{cases} FI_1^{(t)} \left(P^{(1)}(t, x) \right) = \frac{1}{2t^2} \\ FI_1^{(x_k)} \left(P^{(1)}(t, x) \right) = \frac{1}{2x_k^2} \end{cases} \tag{90}$$

$k = 1, 2, \dots, d$.

Then, we can obtain quotients of the squared difference between Jeffreys divergence and space–time variables

$$\begin{cases} \frac{JD(P^{(1)}(t, x), P^{(1)}(s, x))}{|t - s|^2} = \frac{1}{2st} \\ \frac{JD(P^{(1)}(t, x), P^{(1)}(t, \tilde{x}^{(k)}))}{|x_k - x'_k|^2} = \frac{1}{2x_k x'_k} \end{cases} \tag{91}$$

and then we can obtain the relation between quotients and generalized Fisher information

$$\begin{cases} \frac{JD(P^{(1)}(t, x), P^{(1)}(s, x))}{|t - s|^2 FI_1^{(t)}(P^{(1)}(t, x))} = \frac{t}{s} \\ \frac{JD(P^{(1)}(t, x), P^{(1)}(t, \tilde{x}^{(k)}))}{|x_k - x'_k|^2 FI_1^{(x_k)}(P^{(1)}(t, x))} = \frac{x_k}{x'_k} \end{cases} \tag{92}$$

for $k = 1, 2, \dots, d$. If we consider the approximation of spacetime points (t, x) and (s, y) , the final result (92) satisfies the conclusion of Theorem 2.

Similarly, we can obtain the Jeffreys divergence of Brownian sheet bridge at different space–time points

$$JD(P^{(2)}(t, x), P^{(2)}(s, y)) = \frac{\text{prod}(s, y)[1 - \text{prod}(s, y)]}{2\text{prod}(t, x)[1 - \text{prod}(t, x)]} + \frac{\text{prod}(t, x)[1 - \text{prod}(t, x)]}{2\text{prod}(s, y)[1 - \text{prod}(s, y)]} - 1 \tag{93}$$

and the generalized Fisher information on space–time variables

$$\begin{cases} FI_1^{(t)}(P^{(2)}(t, x)) = \frac{[1 - 2\text{prod}(t, x)]^2}{2t^2[1 - \text{prod}(t, x)]^2} \\ FI_1^{(x_k)}(P^{(2)}(t, x)) = \frac{[1 - 2\text{prod}(t, x)]^2}{2x_k^2[1 - \text{prod}(t, x)]^2} \end{cases} \tag{94}$$

$k = 1, 2, \dots, d$. Further, we can easily get the quotients of the squared difference between Jeffreys divergence and space–time variables

$$\begin{cases} \frac{JD(P^{(2)}(t, x), P^{(2)}(s, x))}{|t - s|^2} = \frac{[1 - \text{prod}(x)(s + t)]^2}{2st[1 - \text{prod}(s, x)][1 - \text{prod}(t, x)]} \\ \frac{JD(P^{(2)}(t, x), P^{(2)}(t, \tilde{x}^{(k)}))}{|x_k - x'_k|^2} = \frac{1}{2x_k x'_k [1 - \text{prod}(t, x)][1 - \text{prod}(t, \tilde{x}^{(k)})]} \left[1 - \frac{\text{prod}(t, x)}{x_k} (x_k + x'_k) \right]^2 \end{cases} \tag{95}$$

and then we can obtain the relation between quotients and generalized Fisher information

$$\begin{cases} \frac{JD(P^{(2)}(t, x), P^{(2)}(s, x))}{|t - s|^2 FI_1^{(t)}(P^{(2)}(t, x))} = \frac{t[1 - \text{prod}(t, x)][1 - \text{prod}(x)(s + t)]^2}{s[1 - \text{prod}(s, x)][1 - 2\text{prod}(t, x)]^2} \\ \frac{JD(P^{(2)}(t, x), P^{(2)}(t, \tilde{x}^{(k)}))}{|x_k - x'_k|^2 FI_1^{(x_k)}(P^{(2)}(t, x))} = \frac{x_k[1 - \text{prod}(t, x)]}{x'_k [1 - \text{prod}(t, \tilde{x}^{(k)})][1 - 2\text{prod}(t, x)]^2} \left[1 - \frac{\text{prod}(t, x)}{x_k} (x_k + x'_k) \right]^2 \end{cases} \tag{96}$$

for $k = 1, 2, \dots, d$. Without loss of generality, the result (96) also satisfies Theorem 2.

Next, we evaluate the Jeffreys divergence between the density functions (80) and (85) for the same space–time points. It should be noted that the Brownian sheet bridge density

function is defined on a bounded domain; therefore, we limit our analysis to the space–time region of $(t, x) \in [0, 1] \times [0, 1]^d$.

The Jeffreys divergence between $P^{(1)}$ and $P^{(2)}$ can be easily obtained as

$$JD(P^{(1)}(t, x), P^{(2)}(t, x)) = \frac{1 - \text{prod}(t, x)}{2} + \frac{1}{2[1 - \text{prod}(t, x)]} - 1 \tag{97}$$

and the Fisher divergence as shown in (88) is given by

$$\begin{cases} FD_{(b_0^{(1)}, b_0^{(2)})}(P^{(1)}(t, x) \| P^{(2)}(t, x)) = \text{prod}(x) - \frac{\text{prod}(x)}{[1 - \text{prod}(t, x)]^2} \\ FD_{(b_k^{(1)}, b_k^{(2)})}(P^{(1)}(t, x) \| P^{(2)}(t, x)) = \frac{\text{prod}(t, x)}{x_k} - \frac{\text{prod}(t, x)}{x_k[1 - \text{prod}(t, x)]^2} \end{cases} \tag{98}$$

with the remainder terms

$$R_0(P^{(1)}(t, x) \| P^{(2)}(t, x)) = R_k(P^{(1)}(t, x) \| P^{(2)}(t, x)) = 0 \tag{99}$$

for $k = 1, 2, \dots, d$. Furthermore, we can obtain the generalized De Bruijn identities

$$\begin{cases} \frac{\partial}{\partial t} JD(P^{(1)}(t, x), P^{(2)}(t, x)) = -\frac{1}{2} FD_{(b_0^{(1)}, b_0^{(2)})}(P^{(1)}(t, x) \| P^{(2)}(t, x)) \\ \frac{\partial}{\partial x_k} JD(P^{(1)}(t, x), P^{(2)}(t, x)) = -\frac{1}{2} FD_{(b_k^{(1)}, b_k^{(2)})}(P^{(1)}(t, x) \| P^{(2)}(t, x)) \\ k = 1, 2, \dots, d \end{cases} \tag{100}$$

Next, we present two categories of significant Fokker–Planck equations and provide pertinent illustrations for computing Jeffreys divergence, Fisher information, and Fisher divergence.

Theorem 5. Suppose the functions in the Fokker–Planck Equations (43) for the continuous random field $X(t, x)$ are formulated as follows:

$$\begin{cases} a_k(u; t, x) \equiv 0 \\ b_k(u; t, x) = b_k(t, x)u^2 > 0 \\ k = 0, 1, 2, \dots, d \end{cases} \tag{101}$$

where b_k are continuously differentiable functions independent of u and a continuously differentiable function $\beta(t, x)$ exists, such that

$$d\beta(t, x) = b_0(t, x)dt + b_1(t, x)dx_1 + \dots + b_d(t, x)dx_d \tag{102}$$

the initial value $X(t, x) = 1$ as $\text{prod}(t, x) = 0$ and the initial density function is $p(u; t, x) = \delta(u - 1)$ as $\text{prod}(t, x) = 0$. Then, the density function is

$$p(u; t, x) = \frac{e^{\beta(t, x)}}{\sqrt{2\pi\beta(t, x)}} e^{-\frac{[\log u + \frac{3}{2}\beta(t, x)]^2}{2\beta(t, x)}} \tag{103}$$

Proof. Depending on the conditions, it is easy to obtain the Fokker–Planck equations as

$$\begin{cases} \frac{\partial}{\partial t} p(u; t, x) = \frac{b_0(t, x)}{2} u^2 \frac{\partial^2 p(u; t, x)}{\partial u^2} + 2b_0(t, x)u \frac{\partial}{\partial u} p(u; t, x) + b_0(t, x)p(u; t, x) \\ \frac{\partial}{\partial x_k} p(u; t, x) = \frac{b_k(t, x)}{2} u^2 \frac{\partial^2 p(u; t, x)}{\partial u^2} + 2b_k(t, x)u \frac{\partial}{\partial u} p(u; t, x) + b_k(t, x)p(u; t, x) \\ k = 1, 2, \dots, d \end{cases} \tag{104}$$

Take the transformation $v = \log u$ or $u = e^v$ and note $\tilde{p}(v; t, x) = p(u(v); t, x)$; we can obtain

$$\begin{cases} \frac{\partial}{\partial t} \tilde{p}(v; t, x) = \frac{b_0(t, x)}{2} \frac{\partial^2}{\partial v^2} \tilde{p}(v; t, x) + \frac{3b_0(t, x)}{2} \frac{\partial}{\partial v} \tilde{p}(v; t, x) + b_0(t, x) \tilde{p}(v; t, x) \\ \frac{\partial}{\partial x_k} \tilde{p}(v; t, x) = \frac{b_k(t, x)}{2} \frac{\partial^2}{\partial v^2} \tilde{p}(v; t, x) + \frac{3b_k(t, x)}{2} \frac{\partial}{\partial v} \tilde{p}(v; t, x) + b_k(t, x) \tilde{p}(v; t, x) \\ k = 1, 2, \dots, d \end{cases} \tag{105}$$

with the solution

$$\begin{cases} \tilde{p}(v; t, x) = \frac{e^{\int_0^t b_0(s, x) ds}}{\sqrt{2\pi \int_0^t b_0(s, x) ds}} e^{-\frac{[v + \frac{3}{2} \int_0^t b_0(s, x) ds]^2}{2 \int_0^t b_0(s, x) ds}} \\ \tilde{p}(v; t, x) = \frac{e^{\int_0^{x_k} b_k(t, x) dx_k}}{\sqrt{2\pi \int_0^{x_k} b_k(t, x) dx_k}} e^{-\frac{[v + \frac{3}{2} \int_0^{x_k} b_k(t, x) dx_k]^2}{2 \int_0^{x_k} b_k(t, x) dx_k}} \\ k = 1, 2, \dots, d \end{cases} \tag{106}$$

then,

$$\begin{cases} p(u; t, x) = \frac{e^{\int_0^t b_0(s, x) ds}}{\sqrt{2\pi \int_0^t b_0(s, x) ds}} e^{-\frac{[\log u + \frac{3}{2} \int_0^t b_0(s, x) ds]^2}{2 \int_0^t b_0(s, x) ds}} \\ p(u; t, x) = \frac{e^{\int_0^{x_k} b_k(t, x) dx_k}}{\sqrt{2\pi \int_0^{x_k} b_k(t, x) dx_k}} e^{-\frac{[\log u + \frac{3}{2} \int_0^{x_k} b_k(t, x) dx_k]^2}{2 \int_0^{x_k} b_k(t, x) dx_k}} \\ k = 1, 2, \dots, d \end{cases} \tag{107}$$

Recall that a continuously differential function $\beta(t, x)$ exists such that

$$d\beta(t, x) = b_0(t, x)dt + b_1(t, x)dx_1 + \dots + b_d(t, x)dx_d \tag{108}$$

this enables the derivation of the probability density

$$p(u; t, x) = \frac{e^{\beta(t, x)}}{\sqrt{2\pi\beta(t, x)}} e^{-\frac{[\log u + \frac{3}{2}\beta(t, x)]^2}{2\beta(t, x)}} \tag{109}$$

□

Remark 1. In the stochastic process theory, a correlation exists between the Fokker–Planck equation and the Itô process. Specifically, if the Itô process is

$$dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dB_t \tag{110}$$

then the corresponding Fokker–Planck equation can be obtained as

$$\frac{\partial}{\partial t} p(u, t) = \frac{1}{2} \frac{\partial^2}{\partial u^2} [\sigma^2(u, t)p(u, t)] - \frac{\partial}{\partial u} [\mu(u, t)p(u, t)] \tag{111}$$

where μ and σ represent drift and diffusion, B_t is the standard Brownian motion, or

$$\frac{dX_t}{dt} = \mu(X_t, t) + \sigma(X_t, t)W_t \tag{112}$$

where $W_t = \frac{dB_t}{dt}$ represents the white noise. Actually, if we consider the Itô processes corresponding to Fokker–Planck equations from Theorem 5, we can obtain

$$\begin{cases} \partial_t X(t, x) = \sqrt{b_0(t, x)} X(t, x) W_t \\ \partial_{x_k} X(t, x) = \sqrt{b_k(t, x)} X(t, x) W_k \end{cases} \tag{113}$$

where W_k represents the space white noise with respect to $x_k, k = 1, 2, \dots, d$. Further, we can also write Equation (113) in vector form

$$\nabla X(t, x) = \gamma(t, x) X(t, x) \odot W(t, x) \tag{114}$$

where

$$\begin{cases} \gamma(t, x) = [\sqrt{b_0(t, x)}, \sqrt{b_1(t, x)}, \dots, \sqrt{b_d(t, x)}] \\ W(t, x) = [W_t, W_1, \dots, W_d] \end{cases} \tag{115}$$

∇ represents the gradient operator and \odot represents element by element multiplication. Notice that each equation in Equation (113) is similar in form to the geometric Brownian motion in the theory of stochastic processes. Similarly, we can call the space–time random field that satisfies Equation (113) a geometric Brownian filed.

If we consider different $\beta_3(t, x)$ and $\beta_4(t, x)$ in density function (103), we can obtain density functions $p^{(3)}(u; t, x)$ and $p^{(4)}(u; t, x)$; then, we can obtain the Jeffreys divergence

$$JD(P^{(3)}(t, x), P^{(3)}(s, y)) = \frac{\beta_3(t, x) + \beta_3(s, t) + 4}{8\beta_3(t, x)\beta_3(s, y)} [\beta_3(t, x) - \beta_3(s, y)]^2 \tag{116}$$

and generalized Fisher information

$$\begin{cases} FI_1^{(t)}(P^{(3)}(t, x)) = \frac{\beta_3(t, x) + 2}{4\beta_3^2(t, x)} [b_0^{(3)}(t, x)]^2 \\ FI_1^{(x_k)}(P^{(3)}(t, x)) = \frac{\beta_3(t, x) + 2}{4\beta_3^2(t, x)} [b_k^{(3)}(t, x)]^2 \end{cases} \tag{117}$$

and then the quotients

$$\begin{cases} \frac{JD(P^{(3)}(t, x), P^{(3)}(s, x))}{|t - s|^2} = \frac{\beta_3(t, x) + \beta_3(s, x) + 4}{8\beta_3(t, x)\beta_3(s, y)} \left[\frac{\beta_3(t, x) - \beta_3(s, y)}{t - s} \right]^2 \\ \frac{JD(P^{(3)}(t, x), P^{(3)}(t, \tilde{x}^{(k)}))}{|x_k - x'_k|^2} = \frac{\beta_3(t, x) + \beta_3(t, \tilde{x}^{(k)}) + 4}{8\beta_3(t, x)\beta_3(t, \tilde{x}^{(k)})} \left[\frac{\beta_3(t, x) - \beta_3(t, \tilde{x}^{(k)})}{x_k - x'_k} \right]^2 \end{cases} \tag{118}$$

and we can easily obtain the relation

$$\begin{cases} \frac{JD(P^{(3)}(t, x), P^{(3)}(s, y))}{|t - s|^2 FI_1^{(t)}(P^{(3)}(t, x))} = \frac{\beta_3(t, x)}{\beta_3(s, x)} \frac{\beta_3(t, x) + \beta_3(s, x) + 4}{2[\beta_3(t, x) + 2]} \left[\frac{\beta_3(t, x) - \beta_3(s, y)}{b_0^{(3)}(t, x)(t - s)} \right]^2 \\ \frac{JD(P^{(3)}(t, x), P^{(3)}(t, \tilde{x}^{(k)}))}{|x_k - x'_k|^2 FI_1^{(x_k)}(P^{(3)}(t, x))} = \frac{\beta_3(t, x)}{\beta_3(t, \tilde{x}^{(k)})} \frac{\beta_3(t, x) + \beta_3(t, \tilde{x}^{(k)}) + 4}{2[\beta_3(t, x) + 2]} \left[\frac{\beta_3(t, x) - \beta_3(t, \tilde{x}^{(k)})}{b_k^{(3)}(t, x)(x_k - x'_k)} \right]^2 \end{cases} \tag{119}$$

for $k = 1, 2, \dots, d$. Without a loss of generality, the result (119) also corroborates Theorem 2.

Furthermore, if we consider different $\beta_3(t, x)$ and $\beta_4(t, x)$ in density function (103), we can obtain density functions $p^{(3)}(u; t, x)$ and $p^{(4)}(u; t, x)$; then, the generalized Fisher divergence at the same space–time points is

$$\begin{aligned}
 & FD_{(b_k^{(3)}, b_k^{(4)})} (P^{(3)}(t, x) \| P^{(4)}(t, x)) \\
 &= \frac{\beta_3(t, x) - \beta_4(t, x)}{4\beta_3^2(t, x)\beta_4^2(s, y)} \\
 &\times \left\{ b_k^{(4)}(t, x)\beta_3(t, x) [\beta_3^2(t, x) - 2\beta_4^2(s, y)] - b_k^{(3)}(t, x)\beta_4(t, x) [\beta_4^2(s, y) - 2\beta_3^2(t, x)] \right\} \\
 &\times [4\beta_3(t, x) + 4\beta_4(t, x) - 3\beta_3(t, x)\beta_4(t, x)] [b_k^{(4)}(t, x)\beta_3(t, x) - b_k^{(3)}(t, x)\beta_4(t, x)]
 \end{aligned} \tag{120}$$

with the remainder terms

$$R_k(P^{(3)}(t, x) \| P^{(4)}(t, x)) = 2 [b_k^{(3)}(t, x) - b_k^{(4)}(t, x)] \tag{121}$$

$k = 0, 1, 2, \dots, d$. Then, the generalized De Bruijn identities are as follows:

$$\begin{cases} \frac{\partial}{\partial t} JD(P^{(3)}(t, x), P^{(4)}(t, x)) = -\frac{1}{2} FD_{(b_0^{(3)}, b_0^{(4)})} (P^{(3)}(t, x) \| P^{(4)}(t, x)) \\ \hspace{10em} - 2 [b_0^{(3)}(t, x) - b_0^{(4)}(t, x)] \\ \frac{\partial}{\partial x_k} JD(P^{(3)}(t, x), P^{(4)}(t, x)) = -\frac{1}{2} FD_{(b_k^{(3)}, b_k^{(4)})} (P^{(3)}(t, x) \| P^{(4)}(t, x)) \\ \hspace{10em} - 2 [b_k^{(3)}(t, x) - b_k^{(4)}(t, x)] \\ k = 1, 2, \dots, d \end{cases} \tag{122}$$

Additionally, we offer an alternative non-trivial form below that utilizes the implicit functions method to express our results. This form differs from the one presented in Theorem 5.

Theorem 6. Suppose the functions in Fokker–Planck Equations (43) for the continuously bounded random field $X(t, x) \in [0, 1]$ are formulated as follows:

$$\begin{cases} a_k(u; t, x) = -\frac{3}{2} b_k(t, x)u \\ b_k(u; t, x) = b_k(t, x) (1 - u^2) \\ k = 0, 1, 2, \dots, d \end{cases} \tag{123}$$

where b_k are continuously differentiable functions independent of u and a continuously differentiable function $\beta(t, x)$ exists such that

$$d\beta(t, x) = b_0(t, x)dt + b_1(t, x)dx_1 + \dots + b_d(t, x)dx_d \tag{124}$$

the initial value $X(t, x) = 0$ as $\text{prod}(t, x) = 0$ and the initial density function is $p(u; t, x) = \delta(u)$ as $\text{prod}(t, x) = 0$. Then, the density function is as follows:

$$\begin{cases} p(u; t, x) = \frac{e^{\frac{1}{2}\beta(t, x)}}{\sqrt{2\pi\beta(t, x)}} e^{-\frac{u^2}{2\beta(t, x)}} \\ u = \sin v \end{cases} \tag{125}$$

Proof. Depending on the conditions, it is easy to obtain the Fokker–Planck equations as

$$\begin{cases} \frac{\partial}{\partial t} p(u; t, x) = \frac{b_0(t, x)}{2} \frac{\partial^2}{\partial u^2} [(1 - u^2)p(u; t, x)] - \frac{3b_0(t, x)}{2} \frac{\partial}{\partial u} [up(u; t, x)] \\ \frac{\partial}{\partial x_k} p(u; t, x) = \frac{b_k(t, x)}{2} \frac{\partial^2}{\partial u^2} [(1 - u^2)p(u; t, x)] - \frac{3b_k(t, x)}{2} \frac{\partial}{\partial u} [up(u; t, x)] \\ k = 1, 2, \dots, d \end{cases} \quad (126)$$

By implementing the transformation with $u = \sin v$ and defining $\tilde{p}(v; t, x) = p(u(v; t, x))$, the equations can be restructured as

$$\begin{cases} \frac{\partial}{\partial t} \tilde{p}(v; t, x) = \frac{b_0(t, x)}{2} \frac{\partial^2}{\partial v^2} \tilde{p}(v; t, x) + \frac{b_0(t, x)}{2} \tilde{p}(v; t, x) \\ \frac{\partial}{\partial x_k} \tilde{p}(v; t, x) = \frac{b_k(t, x)}{2} \frac{\partial^2}{\partial v^2} \tilde{p}(v; t, x) + \frac{b_k(t, x)}{2} \tilde{p}(v; t, x) \\ k = 1, 2, \dots, d \end{cases} \quad (127)$$

with the solution

$$\begin{cases} \tilde{p}(v; t, x) = \frac{e^{\frac{1}{2} \int_0^t b_0(s, x) ds}}{\sqrt{2\pi \int_0^t b_0(s, x) ds}} e^{-\frac{v^2}{2 \int_0^t b_0(s, x) ds}} \\ \tilde{p}(v; t, x) = \frac{e^{\frac{1}{2} \int_0^{x_k} b_k(t, x) dx_k}}{\sqrt{2\pi \int_0^{x_k} b_k(t, x) dx_k}} e^{-\frac{v^2}{2 \int_0^{x_k} b_k(t, x) dx_k}} \\ k = 1, 2, \dots, d \end{cases} \quad (128)$$

Recall that a continuously differential function $\beta(t, x)$ exists such that

$$d\beta(t, x) = b_0(t, x)dt + b_1(t, x)dx_1 + \dots + b_d(t, x)dx_d \quad (129)$$

we can derive the probability density function

$$\tilde{p}(v; t, x) = \frac{e^{\frac{1}{2}\beta(t, x)}}{\sqrt{2\pi\beta(t, x)}} e^{-\frac{v^2}{2\beta(t, x)}} \quad (130)$$

then,

$$\begin{cases} p(u; t, x) = \frac{e^{\frac{1}{2}\beta(t, x)}}{\sqrt{2\pi\beta(t, x)}} e^{-\frac{v^2}{2\beta(t, x)}} \\ u = \sin v \end{cases} \quad (131)$$

□

Remark 2. Similar to the discussion in Remark 1, we can obtain the Itô processes corresponding to the Fokker–Planck equations in Theorem 6

$$\begin{cases} \partial_t X(t, x) = -\frac{3}{2} b_0(t, x) X(t, x) + \sqrt{b_0(t, x)[1 - X^2(t, x)]} W_t \\ \partial_{x_k} X(t, x) = -\frac{3}{2} b_k(t, x) X(t, x) + \sqrt{b_k(t, x)[1 - X^2(t, x)]} W_k \end{cases} \quad (132)$$

$k = 1, 2, \dots, d$. In fact, this random field can be solved with a sinusoidal transformation, and the corresponding probability density function can be obtained. Although random field (132) has not yet found its application scenario, it gives us ideas for constructing different forms on space–time random fields in the future.

From density function (125), if we consider different $\beta_5(t, x)$ and $\beta_6(t, x)$, we can obtain density functions $p^{(5)}(u; t, x)$ and $p^{(6)}(u; t, x)$; then, we can obtain the Jeffreys divergence and generalized Fisher information

$$JD\left(P^{(5)}(t, x), P^{(5)}(s, y)\right) = \frac{1 - \beta_5(t, x) - \beta_5(s, y)}{2\beta_5(t, x)\beta_5(s, y)} [\beta_5(t, x) - \beta_5(s, y)]^2 \tag{133}$$

and

$$\begin{cases} FI_1^{(t)}\left(P^{(5)}(t, x)\right) = \frac{1 - 2\beta_5(t, x)}{2\beta_5^2(t, x)} \left[b_0^{(5)}(t, x)\right]^2 \\ FI_1^{(x_k)}\left(P^{(5)}(t, x)\right) = \frac{1 - 2\beta_5(t, x)}{2\beta_5^2(t, x)} \left[b_k^{(5)}(t, x)\right]^2 \end{cases} \tag{134}$$

and then the quotients

$$\begin{cases} \frac{JD\left(P^{(5)}(t, x), P^{(5)}(s, x)\right)}{|t - s|^2} = \frac{1 - \beta_5(t, x) - \beta_5(s, x)}{2\beta_5(t, x)\beta_5(s, x)} \left[\frac{\beta_5(t, x) - \beta_5(s, x)}{t - s}\right]^2 \\ \frac{JD\left(P^{(5)}(t, x), P^{(5)}(t, \tilde{x}^{(k)})\right)}{|x_k - x'_k|^2} = \frac{1 - \beta_5(t, x) - \beta_5(t, \tilde{x}^{(k)})}{2\beta_5(t, x)\beta_5(t, \tilde{x}^{(k)})} \left[\frac{\beta_5(t, x) - \beta_5(t, \tilde{x}^{(k)})}{x_k - x'_k}\right]^2 \end{cases} \tag{135}$$

Obviously, we can easily obtain

$$\begin{cases} \frac{JD\left(P^{(5)}(t, x), P^{(5)}(s, x)\right)}{|t - s|^2 FI_1^{(t)}\left(P^{(5)}(t, x)\right)} = \frac{1 - \beta_5(t, x) - \beta_5(s, x)}{1 - 2\beta_5(t, x)} \frac{\beta_5(t, x)}{\beta_5(s, x)} \left[\frac{\beta_5(t, x) - \beta_5(s, x)}{b_0^{(5)}(t, x)(t - s)}\right]^2 \\ \frac{JD\left(P^{(5)}(t, x), P^{(5)}(t, \tilde{x}^{(k)})\right)}{|x_k - x'_k|^2 FI_1^{(x_k)}\left(P^{(5)}(t, x)\right)} = \frac{1 - \beta_5(t, x) - \beta_5(t, \tilde{x}^{(k)})}{1 - 2\beta_5(t, x)} \frac{\beta_5(t, x)}{\beta_5(t, \tilde{x}^{(k)})} \left[\frac{\beta_5(t, x) - \beta_5(t, \tilde{x}^{(k)})}{b_k^{(5)}(t, x)(x - x'_k)}\right]^2 \end{cases} \tag{136}$$

for $k = 1, 2, \dots, d$. Without a loss of generality, the result (136) corroborates Theorem 2.

Furthermore, if we consider different $\beta_5(t, x)$ and $\beta_6(t, x)$ in the density function (125), denoted as $p^{(5)}(u; t, x)$ and $p^{(6)}(u; t, x)$, we can obtain the generalized Fisher divergence at the same space–time points

$$\begin{aligned} & FD_{(b_k^{(5)}, b_k^{(6)})}\left(P^{(5)}(t, x) \| P^{(6)}(t, x)\right) \\ &= \frac{[\beta_5(t, x) - \beta_6(t, x)] [b_k^{(6)}\beta_5(t, x) - b_k^{(5)}\beta_6(t, x)]}{2\beta_5^2(t, x)\beta_6^2(t, x)} [\beta_5(t, x) - \beta_5^2(t, x) + \beta_6(t, x) - \beta_6^2(t, x)] \end{aligned} \tag{137}$$

with the remainder terms

$$R_k\left(P^{(5)}(t, x) \| P^{(6)}(t, x)\right) = b_k^{(5)}(t, x) - b_k^{(6)}(t, x) \tag{138}$$

for $k = 0, 1, 2, \dots, d$. Then, the generalized De Bruijn identities are as follows:

$$\begin{cases} \frac{\partial}{\partial t} JD\left(P^{(5)}(t, x), P^{(6)}(t, x)\right) = -\frac{1}{2} FD_{(b_0^{(5)}, b_0^{(6)})}\left(P^{(5)}(t, x) \| P^{(6)}(t, x)\right) - [b_0^{(5)}(t, x) - b_0^{(6)}(t, x)] \\ \frac{\partial}{\partial x_k} JD\left(P^{(5)}(t, x), P^{(6)}(t, x)\right) = -\frac{1}{2} FD_{(b_k^{(5)}, b_k^{(6)})}\left(P^{(5)}(t, x) \| P^{(6)}(t, x)\right) - [b_k^{(5)}(t, x) - b_k^{(6)}(t, x)] \end{cases} \tag{139}$$

$k = 1, 2, \dots, d$

5. Conclusions

In this paper, we provide a generalization of the classical definitions of entropy, divergence, and Fisher information and derive these measures on a space–time random field. In addition, the Fokker–Planck Equations (43) for the space–time random field

and density functions are obtained. Moreover, we obtain the Jeffreys divergence of a space–time random field at different space–time positions, and we obtain the approximation of the ratio of Jeffreys divergence to the square of space–time coordinate difference to the generalized Fisher information (54). Additionally, we use the Jeffreys divergence on two space–time random fields from the same type but different parameters Fokker–Planck equations, to obtain generalized De Bruijn identities (61). Finally, we give three examples of Fokker–Planck equations, with their solutions, to calculate the corresponding Jeffreys divergence, generalized Fisher information, and Fisher divergence and obtain the De Bruijn identities. These results encourage further research into the entropy divergence of space–time random fields, which advances the pertinent fields of information entropy, Fisher information, and De Bruijn identities.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The author would like to thank Pingyi Fan, Zhanjie Song, Ying Li, and Yumeng Song for providing relevant references and helpful discussions on topics related to this work.

Conflicts of Interest: The author declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

KL	Kullback–Leibler divergence
FI	Fisher information
CFI	Cross–Fisher information
FD	Fisher divergence
sFD	Symmetric Fisher divergence
JD	Jeffreys divergence

References

1. Risken, H. *The Fokker–Planck Equation: Methods of Solution and Applications*; Springer: Berlin/Heidelberg, Germany, 1984.
2. Stam, A.J. Some inequalities satisfied by the quantities of information of Fisher and Shannon. *Inf. Control* **1959**, *2*, 101–112. [[CrossRef](#)]
3. Barron, A.R. Entropy and the central limit theorem. *Ann. Probab.* **1986**, *14*, 336–342. [[CrossRef](#)]
4. Johnson, O. *Information Theory and the Central Limit Theorem*; Imperial College Press: London, UK, 2004.
5. Guo, D. Relative entropy and score function: New information estimation relationships through arbitrary additive perturbation. In Proceedings of the 2009 IEEE International Symposium on Information Theory, Seoul, Republic of Korea, 28 June–3 July 2009; pp. 814–818.
6. Toranzo, I.V.; Zozor, S.; Brossier, J.-M. Generalization of the De Bruijn Identity to General ϕ -Entropies and ϕ -Fisher Informations. *IEEE Trans. Inform. Theory* **2018**, *64*, 6743–6758. [[CrossRef](#)]
7. Kharazmi, O.; Balakrishnan, N. Cumulative residual and relative cumulative residual Fisher information and their properties. *IEEE Trans. Inform. Theory* **2021**, *67*, 6306–6312. [[CrossRef](#)]
8. Kolmogorov, A.N. The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers. *Dokl. Akad. Nauk SSSR* **1941**, *30*, 299–303.
9. Kolmogorov, A.N. On the degeneration of isotropic turbulence in an incompressible viscous flu. *Dokl. Akad. Nauk SSSR* **1941**, *31*, 538–542.
10. Kolmogorov, A.N. Dissipation of energy in isotropic turbulence. *Dokl. Akad. Nauk SSSR* **1941**, *32*, 19–21.
11. Yaglom, A.M. Some classes of random fields in n -dimensional space, related to stationary random processes. *Theory Probab. Its Appl.* **1957**, *2*, 273–320. [[CrossRef](#)]
12. Yaglom, A.M. *Correlation Theory of Stationary and Related Random Functions. Volume I: Basic Results*; Springer: New York, NY, USA, 1987.
13. Yaglom, A.M. *Correlation Theory of Stationary and Related Random Functions. Volume II: Supplementary Notes and References*; Springer: Berlin, Germany, 1987.
14. Bowditch, A.; Sun, R. The two-dimensional continuum random field Ising model. *Ann. Probab.* **2022**, *50*, 419–454. [[CrossRef](#)]

15. Bailleul, I.; Catellier, R.; Delarue, F. Propagation of chaos for mean field rough differential equations. *Ann. Probab.* **2021**, *49*, 944–996. [[CrossRef](#)]
16. Wu, L.; Samorodnitsky, G. Regularly varying random fields. *Stoch. Process Their Appl.* **2020**, *130*, 4470–4492. [[CrossRef](#)]
17. Koch, E.; Dombry, C.; Robert, C.Y. A central limit theorem for functions of stationary max-stable random fields on \mathbb{R}^d . *Stoch. Process Their Appl.* **2020**, *129*, 3406–3430. [[CrossRef](#)]
18. Ye, Z. On Entropy and ε -Entropy of Random Fields. Ph.D. Dissertation, Cornell University, Ithaca, NY, USA, 1989.
19. Ye, Z.; Berger, T. A new method to estimate the critical distortion of random fields. *IEEE Trans. Inform. Theory* **1992**, *38*, 152–157. [[CrossRef](#)]
20. Ye, Z.; Berger, T. *Information Measures for Discrete Random Fields*, Science Press: Beijing, China; New York, NY, USA, 1998.
21. Ye, Z.; Yang, W. *Random Field: Network Information Theory and Game Theory*, Science Press: Beijing, China, 2023. (In Chinese)
22. Ma, C. Stationary random fields in space and time with rational spectral densities. *IEEE Trans. Inform. Theory* **2007**, *53*, 1019–1029. [[CrossRef](#)]
23. Hairer, M. A theory of regularity structures. *Invent. Math.* **2014**, *198*, 269–504. [[CrossRef](#)]
24. Hairer, M. Solving the KPZ equation. *Ann. Math.* **2013**, *178*, 559–664. [[CrossRef](#)]
25. Kremp, H.; Perkowski, N. Multidimensional SDE with distributional drift and Lévy noise. *Bernoulli* **2022**, *28*, 1757–1783. [[CrossRef](#)]
26. Beeson, R.; Namachchivaya, N.S.; Perkowski, N. Approximation of the filter equation for multiple timescale, correlated, nonlinear systems. *SIAM J. Math. Anal.* **2022**, *54*, 3054–3090. [[CrossRef](#)]
27. Song, Z.; Zhang, J. A note for estimation about average differential entropy of continuous bounded space–time random field. *Chin. J. Electron.* **2022**, *31*, 793–803. [[CrossRef](#)]
28. Kramers, H.A. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* **1940**, *7*, 284–304. [[CrossRef](#)]
29. Moyal, J.E. Stochastic processes and statistical physics. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **1949**, *11*, 150–210. [[CrossRef](#)]
30. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. 623–656. [[CrossRef](#)]
31. Neeser, F.D.; Massey, J.L. Proper complex random processes with applications to information theory. *IEEE Trans. Inform. Theory* **1991**, *39*, 1293–1302. [[CrossRef](#)]
32. Ihara, S. *Information Theory-for Continuous Systems*; World Scientific: Singapore, 1993.
33. Gray, R.M. *Entropy and Information Theory*; Springer: Boston, MA, USA, 2011.
34. Bach, F. Information Theory With Kernel Methods. *IEEE Trans. Inform. Theory* **2023**, *69*, 752–775. [[CrossRef](#)]
35. Kullback, S.; Leibler, R.A. On information and sufficiency, *Ann. Math. Stat.* **1951**, *22*, 79–86. [[CrossRef](#)]
36. Jeffreys, H. An invariant form for the prior probability in estimation problems. *Proc. R. Soc. Lond. A* **1946**, *186*, 453–461.
37. Fuglede, B.; Topsøe, F. Jensen-Shannon divergence and Hilbert space embedding. In Proceedings of the IEEE International Symposium on Information Theory (ISIT), Chicago, IL, USA, 27 June–2 July 2004; Volume 31.
38. Rényi, A. On measures of entropy and information. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*; University of California Press: Berkeley, CA, USA, 1961; Volume 1, pp. 547–561.
39. She R.; Fan P.; Liu X.-Y.; Wang X. Interpretable Generative Adversarial Networks With Exponential Function. *IEEE Trans. Signal Process.* **2021**, *69*, 3854–3867. [[CrossRef](#)]
40. Liu, S.; She, R.; Zhu, Z.; Fan, P. Storage Space Allocation Strategy for Digital Data with Message Importance. *Entropy* **2020**, *22*, 591. [[CrossRef](#)]
41. She, R.; Liu, S.; Fan, P. Attention to the Variation of Probabilistic Events: Information Processing with Message Importance Measure. *Entropy* **2019**, *21*, 439. [[CrossRef](#)]
42. Wan, S.; Lu, J.; Fan, P.; Letaief, K.B. Information Theory in Formation Control: An Error Analysis to Multi-Robot Formation. *Entropy* **2018**, *20*, 618. [[CrossRef](#)]
43. She, R.; Liu, S.; Fan, P. Recognizing Information Feature Variation: Message Importance Transfer Measure and Its Applications in Big Data. *Entropy* **2018**, *20*, 401. [[CrossRef](#)]
44. Nielsen, F. An Elementary Introduction to Information Geometry. *Entropy* **2020**, *22*, 1100. [[CrossRef](#)] [[PubMed](#)]
45. Nielsen, F. On the Jensen–Shannon Symmetrization of Distances Relying on Abstract Means. *Entropy* **2019**, *21*, 485. [[CrossRef](#)]
46. Nielsen, F.; Nock, R. Generalizing skew Jensen divergences and Bregman divergences with comparative convexity. *IEEE Signal Process. Lett.* **2017**, *24*, 1123–1127. [[CrossRef](#)]
47. Furuichi, S.; Minculete, N. Refined Young Inequality and Its Application to Divergences. *Entropy* **2021**, *23*, 514. [[CrossRef](#)]
48. Pinele, J.; Strapasson, J.E.; Costa, S.I. The Fisher-Rao Distance between Multivariate Normal Distributions: Special Cases, Bounds and Applications. *Entropy* **2020**, *22*, 404. [[CrossRef](#)]
49. Reverter, F.; Oller, J.M. Computing the Rao distance for Gamma distributions. *J. Comput. Appl. Math.* **2003**, *157*, 155–167. [[CrossRef](#)]
50. Pawula, R.F. Generalizations and extensions of the Fokker–Planck–Kolmogorov equations. *IEEE Trans. Inform. Theory* **1967**, *13*, 33–41. [[CrossRef](#)]
51. Pawula, R.F. Approximation of the linear Boltzmann equation by the Fokker–Planck equation. *Phys. Rev.* **1967**, *162*, 186–188. [[CrossRef](#)]

52. Khoshnevisan, D.; Shi, Z. Brownian Sheet and Capacity, *Ann. Probab.* **1999**, *27*, 1135–1159. [[CrossRef](#)]
53. Revuz, D.; Yor, M. *Continuous Martingales and Brownian Motion*, 2nd ed.; Springer: New York, NY, USA, 1999.

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.