

Article

Not peer-reviewed version

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

[Jiaxing Zhang](#) *

Posted Date: 24 August 2023

doi: 10.20944/preprints202308.1738.v1

Keywords: Space-time random field; Fokker-Planck equations; Differential entropy; Jeffreys divergence; Fisher information; De Bruijn identities



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

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

Jiaxing Zhang 

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

Abstract: In this paper, we derive Jeffreys divergence, generalized Fisher divergence and corresponding De Bruijn identities on space-time random field. First, we determine the relation between Fisher information on the space-time random field in one of the space-time points and the ratio of Jeffreys divergence on a space-time random field at distinct space-time positions to the square of coordinate difference. In addition, we also find identities between the partial derivative of the Jeffreys divergence and the generalized Fisher divergence with respect to space-time variables, i.e. the De Bruijn identities, between two space-time random fields obtained by different parameters under the same Fokker-Planck equations. At the end of this 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 accompanying De Bruijn identities.

Keywords: space-time random field; Fokker-Planck equations; differential entropy; jeffreys divergence; Fisher information; De Bruijn identities

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, Fokker-Planck equations for random fields, particularly space-time random fields, does not have a clear form so far. 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 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). In information theory and statistics, Jeffreys divergence is often used to measure the distance between predicted and true distributions, but its drawback is that when there is less overlap between the two distributions, the result is infinity. In order to avoid the divergence results, we consider the relationship between Jeffreys divergence and generalized Fisher information for space-time random field with small differences in space-time parameters.

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

1.1. Space-time random field

Random field was first studied by Kolmogorov[8–10] and it was gradually improved by Yaglom[11–13] in the middle of 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) as 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

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

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 first two terms of the Kramers-Moyal expansion. In statistical mechanics, Fokker-Planck equation is usually describes the time evolution of probability density function of the velocity of a particle under the influence of drag forces and random forces, as in famous Brownian motion. And this equation is often used to find the density function of 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 $h(\cdot)$ represent the differential entropy and $p(\cdot)$ is probability density function of X . However, differential entropy is not easy to calculate and seldom exists. There have been related studies on the entropy of stochastic processes and continuous systems[31–34]. If we consider a classical one-dimensional Gaussian channel model

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

here 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 the density of Y_t satisfies the following Fokker-Planck equation

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

Further, by calculating the differential entropy of Y_t and obtaining its derivative to t , we can get

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

where

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

is the Fisher information of Y_t . The equation (8) here is the De Buriijn 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], Rényi divergence[38], etc. These measures have been 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

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

where μ is a measure of x .

2. Notations, Definitions and Propositions

2.1. Notations and Assumptions

In this paper, we use the following 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 .
- Suppose that our density functions $p(u; t, x)$, $q(u; s, y) \in C^2(\mathbb{R}) \times C^1(\mathbb{R}_+ \times \mathbb{R}^d)$, i.e. $p(u; t, x)$ and $q(u; s, y)$ are derivable twice with respect to u and once with respect to (t, x) or (s, y) , respectively.
- In this paper, we denote that only the k -th coordinate differs in the d -dimensional real vectors as \tilde{x}_k and \tilde{y}_k , where the k -th coordinates are x_k and y_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 first and important information quantity is the Kullback-Leibler divergence for random fields. Same as the classical Kullback-Leibler divergence, we can easily get the Definition 1.

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

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

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

$$KL(P(t, x) \| Q(s, y)) \neq KL(Q(s, y) \| P(t, x)) \quad (12)$$

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 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 function $p(u; t, x)$ and $q(u; s, y)$ is defined as

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

Here, we replaced $\|$ by $,$ 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 space-time random field.

Definition 3: The Generalized Fisher information of 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 $f(\cdot)$ are formed as

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

In particular, as $f \equiv 1$, $FI_1(t, x)$ is the Fisher information in usual case. In addition to equation (14), we have other similar forms of generalized Fisher information

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

and

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

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 4: For arbitrary non-negative function $f(\cdot)$, we assume the generalized Fisher information of random variable X

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

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

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

When $f \equiv 1$, $FI_1(X)$ is the Fisher information in usual case.

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

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

and derivative function

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

If p_X , p_Y and p_Z are never vanish,

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

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

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

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

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

Hence, we have

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

with equality only if

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

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

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

Averaging both sides over the distribution of z

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

i.e.

$$(\mu + \lambda)^2 J_f(X + Y) \leq \mu^2 J_f(X) + \lambda^2 J_f(Y) \quad (28)$$

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

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

According to Definition 3, we can get relevant generalized definitions of generalized Fisher information in general.

Definition 5: 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 nonnegative function $f(\cdot)$ is formed as

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

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

Definition 6: 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 formed as

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

In particular, as $f \equiv 1$, $FD_1(P(t, x) \| Q(s, y))$ is the Fisher divergence in usual case.

Obviously, the generalized Fisher divergence of random fields is not a symmetrical divergence. To get the symmetrical formula, we need to generalize (31) in another field.

Definition 7: 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 functions $f(\cdot)$ and $g(\cdot)$ is formed as

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

In particular, as $f \equiv g$, $FD_{(f,f)}(P(t, x) \| Q(s, y))$ is the generalized Fisher divergence for random fields by one function. In general, $FD_{(f,g)}(P(t, x) \| Q(s, y))$ is asymmetric with respect to P and Q , i.e.

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

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

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

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

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

and we can easily get

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

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

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

Notice that

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

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

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

Proposition 8 (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

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

where

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

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

Proposition 9 (Pawula theorem)[50,51]: If the limit on conditional moment of random process $X(t)$

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

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 $n \geq 3$.

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 focus on the case of Fokker-Planck equation.

3. Main Results and Proofs

Lemma 1: Suppose f, p, q are continuous derivable functions defined on \mathbb{R} . Then

$$\int_{\mathbb{R}} f(u) \left[\frac{d \log p(u)}{du} - \frac{d \log q(u)}{du} \right] [p(u) + q(u)] du = - \int_{\mathbb{R}} f'(u) [p(u) - q(u)] du \quad (43)$$

holds true if it makes sense.

Proof:

$$\begin{aligned} & \int_{\mathbb{R}} f(u) \left[\frac{d \log p(u)}{du} - \frac{d \log q(u)}{du} \right] [p(u) + q(u)] du \\ &= \int_{\mathbb{R}} f(u) p(u) \frac{d \log p(u)}{du} du + \int_{\mathbb{R}} f(u) q(u) \frac{d \log p(u)}{du} du \\ & \quad - \int_{\mathbb{R}} f(u) p(u) \frac{d \log q(u)}{du} du - \int_{\mathbb{R}} f(u) q(u) \frac{d \log q(u)}{du} du \\ &= \int_{\mathbb{R}} f(u) [p'(u) - q'(u)] du + \int_{\mathbb{R}} f(u) \left[\frac{q(u)}{p(u)} p'(u) - \frac{p(u)}{q(u)} q'(u) \right] du \end{aligned} \quad (44)$$

Notice that

$$\begin{cases} \frac{d}{du} \left[\frac{p(u)}{q(u)} \right] = \frac{1}{q(u)} \left[p'(u) - \frac{p(u)}{q(u)} q'(u) \right] \\ \frac{d}{du} \left[\frac{q(u)}{p(u)} \right] = \frac{1}{p(u)} \left[q'(u) - \frac{q(u)}{p(u)} p'(u) \right] \end{cases} \quad (45)$$

i.e.

$$\begin{cases} \frac{p(u)}{q(u)} q'(u) = p'(u) - q(u) \frac{d}{du} \left[\frac{p(u)}{q(u)} \right] \\ \frac{q(u)}{p(u)} p'(u) = q'(u) - p(u) \frac{d}{du} \left[\frac{q(u)}{p(u)} \right] \end{cases} \quad (46)$$

then

$$\begin{aligned} & \int_{\mathbb{R}} f(u) \left[\frac{q(u)}{p(u)} p'(u) - \frac{p(u)}{q(u)} q'(u) \right] du \\ &= \int_{\mathbb{R}} f(u) \left[p'(u) - q(u) \frac{d}{du} \left[\frac{p(u)}{q(u)} \right] \right] du - \int_{\mathbb{R}} f(u) \left[q'(u) - p(u) \frac{d}{du} \left[\frac{q(u)}{p(u)} \right] \right] du \\ &= \int_{\mathbb{R}} f(u) [p'(u) - q'(u)] du - \int_{\mathbb{R}} f(u) q(u) d \left[\frac{p(u)}{q(u)} \right] + \int_{\mathbb{R}} f(u) p(u) d \left[\frac{q(u)}{p(u)} \right] \\ &= \int_{\mathbb{R}} f(u) [p'(u) - q'(u)] du + \int_{\mathbb{R}} [f'(u) q(u) + f(u) q'(u)] \frac{p(u)}{q(u)} du \\ & \quad - \int_{\mathbb{R}} [f'(u) p(u) + f(u) p'(u)] \frac{q(u)}{p(u)} du \\ &= \int_{\mathbb{R}} f(u) [p'(u) - q'(u)] du + \int_{\mathbb{R}} f'(u) [q(u) - p(u)] du - \int_{\mathbb{R}} f(u) \left[\frac{q(u)}{p(u)} p'(u) - \frac{p(u)}{q(u)} q'(u) \right] du \\ &= \int_{\mathbb{R}} d \{ f(u) [q(u) - p(u)] \} - \int_{\mathbb{R}} f(u) \left[\frac{q(u)}{p(u)} p'(u) - \frac{p(u)}{q(u)} q'(u) \right] du \\ &= - \int_{\mathbb{R}} f(u) \left[\frac{q(u)}{p(u)} p'(u) - \frac{p(u)}{q(u)} q'(u) \right] du \end{aligned} \quad (47)$$

So we can get the result

$$\int_{\mathbb{R}} f(u) \left[\frac{d \log p(u)}{du} - \frac{d \log q(u)}{du} \right] [p(u) + q(u)] du = - \int_{\mathbb{R}} f'(u) [p(u) - q(u)] du \quad (48)$$

Lemma 2: Suppose that the Fokker-Planck equation for the density function $p(u, t)$ is

$$\frac{\partial p(u, t)}{\partial t} = \frac{1}{2} \frac{\partial^2 [b(u, t)p(u, t)]}{\partial u^2} - \frac{\partial [a(u, t)p(u, t)]}{\partial u} \quad (49)$$

We use variable substitution $u = u(v)$ or $v = v(u)$ to convert the equation to

$$\frac{\partial p(v, t)}{\partial t} = \frac{E(t)}{2} \frac{\partial^2 p(v, t)}{\partial v^2} - F(t) \frac{\partial p(v, t)}{\partial v} + G(t)p(v, t) \quad (50)$$

where $E(t)$, $F(t)$, and $G(t)$ are functions of t and $E(t) > 0$. Then we get the density

$$p(u, t) = \frac{e^{\int_0^t G(s)ds}}{\sqrt{2\pi \int_0^t E(s)ds}} e^{-\frac{[v(u)-v(u_0)-\int_0^t F(s)ds]^2}{2 \int_0^t E(s)ds}} \quad (51)$$

Proof: Using Fourier transform, we can get the solution of Fokker-Planck equation (50) about v

$$p(v, t) = \frac{e^{\int_0^t G(s)ds}}{\sqrt{2\pi \int_0^t E(s)ds}} e^{-\frac{[v-v_0(x)-\int_0^t F(s)ds]^2}{2 \int_0^t E(s)ds}} \quad (52)$$

It is worth noting that while p is a probability density function for u , and p is not a probability density function for v . Our transformation, $u = u(v)$ or $v = v(u)$, is only introduced in the process of solving the equation, which does not mean that v is a random process. Therefore, the integral value of p with respect to v is not 1. But we can obtain

$$p(u, t) = \frac{e^{\int_0^t G(s)ds}}{\sqrt{2\pi \int_0^t E(s)ds}} e^{-\frac{[v(u)-v(u_0)-\int_0^t F(s)ds]^2}{2 \int_0^t E(s)ds}} \quad (53)$$

and

$$\int_{I_u} \frac{e^{\int_0^t G(s)ds}}{\sqrt{2\pi \int_0^t E(s)ds}} e^{-\frac{[v(u)-v(u_0)-\int_0^t F(s)ds]^2}{2 \int_0^t E(s)ds}} du = 1 \quad (54)$$

where I_u is the supporting set of $p(u, t)$ with respect to u .

Theorem 3: The probability density function $p(u; t, x)$ of 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

$$\begin{cases} \frac{\partial}{\partial t} p(u; t, x) = \frac{1}{2} \frac{\partial^2}{\partial u^2} [b_0(u; t, x)p(u, t, x)] - \frac{\partial}{\partial u} [a_0(u; t, x)p(u; t, x)] \\ \frac{\partial}{\partial x_k} p(u; t, x) = \frac{1}{2} \frac{\partial^2}{\partial u^2} [b_k(u; t, x)p(u, t, x)] - \frac{\partial}{\partial u} [a_k(u; t, x)p(u; t, x)] \\ k = 1, 2, \dots, d \end{cases} \quad (55)$$

where

$$\begin{cases} a_0(u; t, x) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} M_1(u; t, \Delta t, x) \\ b_0(u; t, x) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} M_2(u; t, \Delta t, x) \\ a_k(u; t, x) = \lim_{\Delta x_k \rightarrow 0} \frac{1}{\Delta x_k} M_1(u; t, x, \Delta x_k) \\ b_k(u; t, x) = \lim_{\Delta x_k \rightarrow 0} \frac{1}{\Delta x_k} M_2(u; t, x, \Delta x_k) \\ k = 1, 2, \dots, d \end{cases} \quad (56)$$

here

$$\begin{cases} M_n(u; t, \Delta t, x) = E \left\{ [X(t + \Delta t, x) - X(t, x)]^n \mid X(t, x) = u \right\} \\ M_n(u; t, x, \Delta x_k) = E \left\{ [X(t, x + \Delta x_k) - X(t, x)]^n \mid X(t, x) = u \right\} \end{cases} \quad (57)$$

are n -order conditional moments and (e_1, e_2, \dots, e_d) is the orthonormal basis of the standard identity of \mathbb{R}^d .

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

$$p(u; t + \Delta t, x) - p(u; t, x) = \sum_{n=1}^{+\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial u^n} [M_n(u; t, \Delta t, x) p(u; t, x)] \quad (58)$$

where

$$M_n(u; t, \Delta t, x) = E \left\{ [X(t + \Delta t, x) - X(t, x)]^n \mid X(t, x) = u \right\} \quad (59)$$

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

$$\frac{\partial}{\partial t} p(u; t, x) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \sum_{n=1}^{+\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial u^n} [M_n(u; t, \Delta t, x) p(u; t, x)] \quad (60)$$

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

$$\frac{\partial}{\partial t} p(u; t, x) = \frac{1}{2} \frac{\partial^2}{\partial u^2} [b_0(u; t, x) p(u, t, x)] - \frac{\partial}{\partial u} [a_0(u; t, x) p(u; t, x)] \quad (61)$$

where

$$\begin{cases} a_0(u; t, x) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} M_1(u; t, \Delta t, x) \\ b_0(u; t, x) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} M_2(u; t, \Delta t, x) \end{cases} \quad (62)$$

Similarly, we may consider the increment Δx_k of the space variable x_k and we can obtain the Fokker-Planck equations about x_k is

$$\frac{\partial}{\partial x_k} p(u; t, x) = \frac{1}{2} \frac{\partial^2}{\partial u^2} [b_k(u; t, x) p(u, t, x)] - \frac{\partial}{\partial u} [a_k(u; t, x) p(u; t, x)] \quad (63)$$

where

$$\begin{cases} a_k(u; t, x) = \lim_{\Delta x_k \rightarrow 0} \frac{1}{\Delta x_k} M_1(u; t, x, \Delta x_k) \\ b_k(u; t, x) = \lim_{\Delta x_k \rightarrow 0} \frac{1}{\Delta x_k} M_2(u; t, x, \Delta x_k) \end{cases} \quad (64)$$

here

$$M_n(u; t, x, \Delta x_k) = E \left\{ [X(t, x + \Delta x_k e_k) - X(t, x)]^n \mid X(t, x) = u \right\} \quad (65)$$

(e_1, e_2, \dots, e_d) is the orthonormal basis of the standard identity of \mathbb{R}^d , $k = 1, 2, \dots, d$.

Theorem 4: Suppose that $p(u; t, x)$ and $p(u; s, y)$ are rapidly decreasing functions of space-time random field $X(t, x)$ such that the partial integral terms are 0, $(t, x), (s, y) \in \mathbb{R}_+ \times \mathbb{R}^d$. If there are two infinitesimal $\varepsilon(t, x, s, y) > 0$ and $\eta(t, x, s, y) > 0$ satisfied

$$\begin{cases} \lim_{|t-s| \rightarrow 0} \varepsilon(t, x, s, y) = \lim_{|t-s| \rightarrow 0} \eta(t, x, s, y) = 0 \\ \lim_{|x_k - y_k| \rightarrow 0} \varepsilon(t, x, s, y) = \lim_{|x_k - y_k| \rightarrow 0} \eta(t, x, s, y) = 0 \\ k = 1, 2, \dots, d \end{cases} \quad (66)$$

where x_k, y_k are the k -th coordinate of $x, y \in \mathbb{R}$, such that

$$1 - \varepsilon(t, x, s, y) < \frac{p(u; t, x)}{p(u; s, y)} < 1 + \eta(t, x, s, y) \quad (67)$$

is true to all meaningful $u \in \mathbb{R}$. Then we have

$$\begin{cases} \frac{1}{1 + \eta(t, x, s, y)} + \lambda(t - s) < \frac{JD(P(t, x), P(s, x))}{|t - s|^2 FI_1^{(s)}(X(s, x))} < \frac{1}{1 - \varepsilon(t, x, s, y)} + \lambda(t - s) \\ \frac{1}{1 + \eta(t, x, s, y)} + \mu_k(x_k - y_k) < \frac{JD(P(t, \tilde{x}_k), P(t, \tilde{y}_k))}{|x_k - y_k|^2 FI_1^{(y_k)}(X(t, \tilde{y}_k))} < \frac{1}{1 - \varepsilon(t, x, s, y)} + \mu_k(x_k - y_k) \end{cases} \quad (68)$$

here $FI_1^{(s)}$ and $FI_1^{(y_k)}$ are generalized Fisher information on space-time variables and

$$\begin{cases} \lim_{|t-s| \rightarrow 0} \lambda(t - s) = 0 \\ \lim_{|x_k - y_k| \rightarrow 0} \mu_k(x_k - y_k) = 0 \end{cases} \quad (69)$$

are infinitesimal on space-time variables, $k = 1, 2, \dots, d$.

Proof: First, we consider $JD(P(t, x), P(s, x)) / |t - s|^2$. Notice that

$$\begin{aligned} & \log p(u; t, x) - \log p(u; s, x) \\ &= \frac{1}{\gamma p(u; t, x) + (1 - \gamma)p(u; s, x)} [p(u; t, x) - p(u; s, x)] \\ &= \frac{1}{\gamma p(u; t, x) + (1 - \gamma)p(u; s, x)} [\partial_s p(u; s, x)(t - s) + o(t - s)] \end{aligned} \quad (70)$$

then we can get

$$\begin{aligned} & \frac{JD(P(t, x), P(s, x))}{|t - s|^2} \\ &= \frac{1}{|t - s|^2} \int_{\mathbb{R}} [\partial_s p(u; s, x)(t - s) + o(t - s)]^2 \frac{1}{\gamma p(u; t, x) + (1 - \gamma)p(u; s, x)} du \\ &= \frac{1}{|t - s|^2} \int_{\mathbb{R}} \left[\partial_s [\log p(u; s, x)] (t - s) + \frac{1}{p(u; s, x)} o(t - s) \right]^2 p(u; s, x) \\ & \quad \times \frac{p(u; s, x)}{\gamma p(u; t, x) + (1 - \gamma)p(u; s, x)} du \\ &= \int_{\mathbb{R}} \left[\partial_s [\log p(u; s, x)] + \frac{1}{p(u; s, x)} \frac{o(t - s)}{|t - s|} \right]^2 p(u; s, x) \frac{1}{\gamma \frac{p(u; t, x)}{p(u; s, x)} + (1 - \gamma)} du \end{aligned} \quad (71)$$

where $0 < \gamma < 1$.

Recall the conditions, there are two infinitesimal $\varepsilon(t, x, s, y)$, $\eta(t, x, s, y) > 0$ and

$$\lim_{|t-s| \rightarrow 0} \varepsilon(t, x, s, y) = \lim_{|t-s| \rightarrow 0} \eta(t, x, s, y) = 0 \quad (72)$$

such that

$$1 - \varepsilon(t, x, s, y) < \frac{p(u; t, x)}{p(u; s, y)} < 1 + \eta(t, x, s, y) \quad (73)$$

is true to all meaningful $u \in \mathbb{R}$, then we can get

$$\begin{aligned} & \frac{1}{1 + \eta(t, x, s, y)} \\ & < \frac{1}{\gamma(1 + \eta(t, x, s, y)) + (1 - \gamma)} \\ & < \frac{1}{\gamma \frac{p(u; t, x)}{p(u; s, y)} + (1 - \gamma)} \\ & < \frac{1}{\gamma(1 - \varepsilon(t, x, s, y)) + (1 - \gamma)} \\ & < \frac{1}{1 - \varepsilon(t, x, s, y)} \end{aligned} \quad (74)$$

so there exists infinitesimal $\lambda(t - s)$ as $|t - s| \rightarrow 0$ such that

$$\frac{1}{1 + \eta(t, x, s, y)} + \lambda(t - s) < \frac{JD(P(t, x), P(s, x))}{|t - s|^2 FI_1^{(s)}(X(s, x))} < \frac{1}{1 - \varepsilon(t, x, s, y)} + \lambda(t - s) \quad (75)$$

Similarly, we can get the bounds of difference quotient on Jeffreys divergence for space coordinates

$$\frac{1}{1 + \eta(t, x, s, y)} + \mu_k(x_k - y_k) < \frac{JD(P(t, \tilde{x}_k), P(t, \tilde{y}_k))}{|x_k - y_k|^2 FI_1^{(y_k)}(X(t, \tilde{y}_k))} < \frac{1}{1 - \varepsilon(t, x, s, y)} + \mu_k(x_k - y_k) \quad (76)$$

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

Theorem 5: Suppose that $p(u; t, x)$ and $q(u; t, x)$ are rapidly decreasing functions of space-time random fields $X(t, x)$ and $Y(t, x)$ such that the partial integral terms are 0, $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^d$. Then the Jeffreys divergence $JD(P(t, x), Q(t, x))$ satisfies generalized De Bruijn identities

$$\begin{cases} \frac{\partial}{\partial t} JD(P(t, x), Q(t, x)) = -\frac{1}{2} FD_{(b_0^{(1)}, b_0^{(2)})}(P(t, x) \| Q(t, x)) - R_0(P(t, x) \| Q(t, x)) \\ \frac{\partial}{\partial x_k} JD(P(t, x), Q(t, x)) = -\frac{1}{2} FD_{(b_k^{(1)}, b_k^{(2)})}(P(t, x) \| Q(t, x)) - R_k(P(t, x) \| Q(t, x)) \\ k = 1, 2, \dots, d \end{cases} \quad (77)$$

where a_0, b_0, a_k, b_k are the forms in (56) and (57), and

$$\begin{cases} R_0(P(t, x) \| Q(t, x)) = \int_{\mathbb{R}} \left[\frac{1}{2} \partial_{uu}^2 (b_0^{(1)} - b_0^{(2)}) - \partial_u (a_0^{(1)} - a_0^{(2)}) \right] (p + q) du \\ R_k(P(t, x) \| Q(t, x)) = \int_{\mathbb{R}} \left[\frac{1}{2} \partial_{uu}^2 (b_k^{(1)} - b_k^{(2)}) - \partial_u (a_k^{(1)} - a_k^{(2)}) \right] (p + q) du \\ k = 1, 2, \dots, d \end{cases} \quad (78)$$

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

Proof: Notice that

$$\begin{aligned}
 JD(P(t, x), Q(t, x)) &= KL(P(t, x) \| Q(t, x)) + KL(Q(t, x) \| P(t, x)) \\
 &= \int_{\mathbb{R}} p \log \frac{p}{q} du + \int_{\mathbb{R}} q \log \frac{q}{p} du \\
 &= \int_{\mathbb{R}} \left[p \log \frac{p}{q} + q \log \frac{q}{p} \right] du du
 \end{aligned} \tag{79}$$

where $p := p(u; t, x)$, $q := q(u; t, x)$ and we omit the variable $(u; t, x)$. 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\} (\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{80}$$

where

$$\begin{aligned}
 &FD_{(b_0^{(1)}, b_0^{(2)})} (P(t, x) \| Q(t, x)) \\
 &= \int_{\mathbb{R}} \left[b_0^{(1)} \partial_u \log p - b_0^{(2)} \partial_u \log q \right] (\partial_u \log p - \partial_u \log q) (p + q) du
 \end{aligned} \tag{81}$$

and

$$R_0 (P(t, x) \| Q(t, x)) = \int_{\mathbb{R}} \left[\frac{1}{2} \partial_{uu}^2 (b_0^{(1)} - b_0^{(2)}) - \partial_u (a_0^{(1)} - a_0^{(2)}) \right] (p + q) du \tag{82}$$

Similarly, for $k = 1, 2, \dots, d$, the generalized De Bruijn identities about the space variable x_k is

$$\frac{\partial}{\partial x_k} JD(P(t, x), Q(t, x)) = -\frac{1}{2} FD_{(b_k^{(1)}, b_k^{(2)})}(P(t, x) \| Q(t, x)) - R_k(P(t, x), Q(t, x)) \quad (83)$$

where

$$R_k(P(t, x) \| Q(t, x)) = \int_{\mathbb{R}} \left[\frac{1}{2} \partial_{uu}^2 (b_k^{(1)} - b_k^{(2)}) - \partial_u (a_k^{(1)} - a_k^{(2)}) \right] (p + q) du \quad (84)$$

then we get the conclusion.

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

In this section, we list three types Fokker-Planck equations, obtain their density functions and the corresponding information measures, the Jeffreys divergence, generalized Fisher information and Fisher divergence. Starting from these quantities, we get two results illustrated in **Theorem 4** and **Theorem 5**. On the one hand, we obtain the quotient of Jeffreys divergence and the squared of space-time variation on the same Fokker-Planck space-time random field at different space-time points, comparing with the generalized Fisher information. On the other hand, we obtain the De Bruijn identities on Jeffreys divergence and generalized Fisher divergence from the Fokker-Planck equations on space-time random field at the same space-time position under the same type and different parameters.

4.1. A Trivial Equation

If we let

$$\begin{cases} a_0(u; t, x) = a_0(t, x) \\ b_0(u; t, x) = b_0(t, x) > 0 \\ a_k(u; t, x) = a_k(t, x) \\ b_k(u; t, x) = b_k(t, x) > 0 \\ k = 1, 2, \dots, d \end{cases} \quad (85)$$

are continuous derivable functions independent of u and $p(u; 0, x) = \delta[u - u_0(x)]$ is the initial density, the Fokker-Planck equations are simple parabolic equations and the solution can be obtained by Fourier transform

$$\begin{cases} p(u; t, x) = \frac{1}{\sqrt{2\pi \int_0^t b_0(s, x) ds}} e^{-\frac{[u - u_0(x) - \int_0^t a_0(s, x) ds]^2}{2 \int_0^t b_0(s, x) ds}} \\ p(u; t, x) = \frac{1}{\sqrt{2\pi \int_0^{x_k} b_k(t, x) dx_k}} e^{-\frac{[u - u_0(x) - \int_0^{x_k} a_k(t, x) dx_k]^2}{2 \int_0^{x_k} b_k(t, x) dx_k}} \end{cases} \quad (86)$$

In order to write the above results into a unified form, we suppose that

$$\begin{cases} \int_0^t a_0(s, x) ds = \int_0^{x_k} a_k(t, x) dx_k \\ \int_0^t b_0(s, x) ds = \int_0^{x_k} b_k(t, x) dx_k \end{cases} \quad (87)$$

hold for $k = 1, 2, \dots, d$, we can get a unified formula. This reminds us that we need to find functions $\alpha(t, x)$ and $\beta(t, x) > 0$ subject to the total differential

$$\begin{cases} d\alpha(t, x) = a_0 dt + a_1 dx_1 + \dots + a_d dx_d \\ d\beta(t, x) = b_0 dt + b_1 dx_1 + \dots + b_d dx_d \end{cases} \quad (88)$$

i.e. $a_0, a_1, \dots, a_d, b_0, b_1, \dots, b_d$ correspond to partial derivatives of $\alpha(t, x)$ and $\beta(t, x)$ with respect to space-time variables, respectively. In this way, we can get the probability density function

$$p(u; t, x) = \frac{1}{\sqrt{2\pi\beta(t, x)}} e^{-\frac{[u - u_0(x) - \alpha(t, x)]^2}{2\beta(t, x)}} \quad (89)$$

Actually, there are many examples whose Fokker-Planck equations fit this form. Let $B(t, x)$ is $(1 + d, 1)$ Brownian sheet. That is, a centered continuous Gaussian process which is indexed by $(1 + d)$ real, positive parameters and takes its values in \mathbb{R} [52,53]. Moreover, its covariance structure is given by

$$E[B(t, x)B(s, y)] = (t \wedge s) \prod_{k=1}^d (x_k \wedge y_k) \quad (90)$$

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 get

$$E[B^2(t, x)] = \text{prod}(t, x) \quad (91)$$

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

$$p^{(1)}(u; t, x) = \frac{1}{\sqrt{2\pi\text{prod}(t, x)}} e^{-\frac{u^2}{2\text{prod}(t, x)}} \quad (92)$$

Moreover, the Fokker-Planck equations are

$$\begin{cases} \frac{\partial}{\partial t} p^{(1)}(u; t, x) = \frac{\text{prod}(x)}{2} \frac{\partial^2}{\partial u^2} p^{(1)}(u, t, x) \\ \frac{\partial}{\partial x_k} p^{(1)}(u; t, x) = \frac{\text{prod}(t, x)}{2x_k} \frac{\partial^2}{\partial u^2} p^{(1)}(u, t, x) \\ k = 1, 2, \dots, d \end{cases} \quad (93)$$

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

Following the construct idea of Brownian bridge on Brownian motion[53], we name

$$B^*(t, x) = B(t, x) - \text{prod}(t, x)B(1, 1, \dots, 1) \quad (94)$$

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

$$E[B^*(t, x)B^*(s, y)] = E[B(t, x)B(s, y)] - \text{prod}(t, x)\text{prod}(s, y) \quad (95)$$

we can get

$$E[B^2(t, x)] = \text{prod}(t, x) [1 - \text{prod}(t, x)] \quad (96)$$

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

$$p^{(2)}(u; t, x) = \frac{1}{\sqrt{2\pi\text{prod}(t, x) [1 - \text{prod}(t, x)]}} e^{-\frac{u^2}{2\text{prod}(t, x)[1 - \text{prod}(t, x)]}} \quad (97)$$

On the other hand, the Fokker-Planck equations are

$$\begin{cases} \frac{\partial}{\partial t} p^{(2)}(u; t, x) = \frac{\text{prod}(x) [1 - 2\text{prod}(t, x)]}{2} \frac{\partial^2}{\partial u^2} p^{(2)}(u, t, x) \\ \frac{\partial}{\partial x_k} p^{(2)}(u; t, x) = \frac{\text{prod}(t, x)}{2x_k} [1 - 2\text{prod}(t, x)] \frac{\partial^2}{\partial u^2} p^{(2)}(u, t, x) \\ k = 1, 2, \dots, d \end{cases} \quad (98)$$

with the initial condition $p(u; t, x) = \delta(u)$ as $\text{prod}(t, x) = 0$ and we get the solution (97).

Now, we have density functions (92) and (97) to get their respective Jeffreys divergence and the generalized De Buriijn identities. Actually, we can easily get the Jeffreys divergence of (89) at different space-time points is formed as

$$JD(P(t, x), P(s, y)) = \frac{[\alpha(t, x) - \alpha(s, y)]^2 + \beta(s, y)}{2\beta(t, x)} + \frac{[\alpha(t, x) - \alpha(s, y)]^2 + \beta(t, x)}{2\beta(s, y)} - 1 \quad (99)$$

and the Fisher divergence of $P^{(1)}$ and $P^{(2)}$ at the same space-time points

$$\begin{aligned} & FD_{(b_k^{(1)}, b_k^{(2)})} (P^{(1)}(t, x) \| P^{(2)}(t, x)) \\ &= \frac{1}{\beta_1^2(t, x) \beta_2^2(t, x)} \left\{ [\alpha_1(t, x) - \alpha_2(t, x)]^2 [b_k^{(2)} \beta_1^2(t, x) + b_k^{(1)} \beta_2^2(t, x)] \right. \\ &\quad \left. + [\beta_1(t, x) - \beta_2(t, x)] [\beta_1(t, x) + \beta_2(t, x)] [b_k^{(2)} \beta_1(t, x) - b_k^{(1)} \beta_2(t, x)] \right\} \end{aligned} \quad (100)$$

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

Bring the density function of Brownian sheet into (99), we can get the Jeffreys divergence of Brownian sheet at different space-time points

$$JD(P^{(1)}(t, x), P^{(1)}(s, y)) = \frac{\text{prod}(s, y)}{2\text{prod}(t, x)} + \frac{\text{prod}(t, x)}{2\text{prod}(s, y)} - 1 \quad (101)$$

and the generalized Fisher information on space-time variables are

$$\begin{cases} FI_1^{(t)}(P^{(1)}(t, x)) = \frac{1}{2t^2} \\ FI_1^{(x_k)}(P^{(1)}(t, x)) = \frac{1}{2x_k^2} \end{cases} \quad (102)$$

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

Then we can get quotients of the squared difference between Jeffreys divergence and space-time variables

$$\begin{cases} \frac{JD(P^{(1)}(t, x), P^{(1)}(s, x))}{|t - s|^2} = \frac{1}{2st} \\ \frac{JD(P^{(1)}(t, \tilde{x}_k), P^{(1)}(t, \tilde{y}_k))}{|x_k - y_k|^2} = \frac{1}{2x_k y_k} \end{cases} \quad (103)$$

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

$$\begin{cases} \frac{JD \left(P^{(1)}(t, x), P^{(1)}(s, x) \right)}{|t - s|^2 FI_1^{(t)} \left(P^{(1)}(t, x) \right)} = \frac{t}{s} \\ \frac{JD \left(P^{(1)}(t, \tilde{x}_k), P^{(1)}(t, \tilde{y}_k) \right)}{|x_k - y_k|^2 FI_1^{(x_k)} \left(P^{(1)}(t, x) \right)} = \frac{x_k}{y_k} \end{cases} \quad (104)$$

for $k = 1, 2, \dots, d$. If we consider the approximation of spacetime points (t, x) and (s, y) , the final result (104) satisfies the conclusion of **Theorem 4**.

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

$$JD \left(P^{(2)}(t, x), P^{(2)}(s, y) \right) = \frac{\text{prod}(s, y) [1 - \text{prod}(s, y)]}{2 \text{prod}(t, x) [1 - \text{prod}(t, x)]} + \frac{\text{prod}(t, x) [1 - \text{prod}(t, x)]}{2 \text{prod}(s, y) [1 - \text{prod}(s, y)]} - 1 \quad (105)$$

and the generalized Fisher information on space-time variables

$$\begin{cases} FI_1^{(t)} \left(P^{(2)}(t, x) \right) = \frac{[1 - 2 \text{prod}(t, x)]^2}{2t^2 [1 - \text{prod}(t, x)]^2} \\ FI_1^{(x_k)} \left(P^{(2)}(t, x) \right) = \frac{[1 - 2 \text{prod}(t, x)]^2}{2x_k^2 [1 - \text{prod}(t, x)]^2} \end{cases} \quad (106)$$

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

$$\begin{cases} \frac{JD \left(P^{(2)}(t, x), P^{(2)}(s, x) \right)}{|t - s|^2} = \frac{[1 - \text{prod}(x)(s + t)]^2}{2st [1 - \text{prod}(s, x)] [1 - \text{prod}(t, x)]} \\ \frac{JD \left(P^{(2)}(t, \tilde{x}_k), P^{(2)}(t, \tilde{y}_k) \right)}{|x_k - y_k|^2} = \frac{1}{2x_k y_k [1 - \text{prod}(t, x)] [1 - \text{prod}(t, y)]} \left[1 - \frac{\text{prod}(t, \tilde{x}_k)}{x_k} (x_k + y_k) \right]^2 \end{cases} \quad (107)$$

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

$$\begin{cases} \frac{JD \left(P^{(2)}(t, x), P^{(2)}(s, x) \right)}{|t - s|^2 FI_1^{(t)} \left(P^{(2)}(t, x) \right)} = \frac{t [1 - \text{prod}(t, x)] [1 - \text{prod}(x)(s + t)]^2}{s [1 - \text{prod}(s, x)] [1 - 2 \text{prod}(t, x)]^2} \\ \frac{JD \left(P^{(2)}(t, \tilde{x}_k), P^{(2)}(t, \tilde{y}_k) \right)}{|x_k - y_k|^2 FI_1^{(x_k)} \left(P^{(2)}(t, x) \right)} = \frac{x_k [1 - \text{prod}(t, \tilde{x}_k)]}{y_k [1 - \text{prod}(t, \tilde{y}_k)] [1 - 2 \text{prod}(t, x)]^2} \left[1 - \frac{\text{prod}(t, \tilde{x}_k)}{x_k} (x_k + y_k) \right]^2 \end{cases} \quad (108)$$

for $k = 1, 2, \dots, d$. Without loss of generality, the result (108) also satisfies **Theorem 4**.

Next, we consider the Jeffreys divergence on (92) and (97) at the same space-time points. Notice that the bounded domain of the Brownian sheet bridge density function, we consider only the space-time area $(t, x) \in [0, 1] \times [0, 1]^d$.

We can easily get the Jeffreys divergence

$$JD \left(P^{(1)}(t, x), P^{(2)}(t, x) \right) = \frac{1 - \text{prod}(t, x)}{2} + \frac{1}{2 [1 - \text{prod}(t, x)]} - 1 \quad (109)$$

and the Fisher divergence from (100)

$$\begin{cases} FD_{(b_0^{(1)}, b_0^{(2)})} (P^{(1)}(t, x) \| P^{(2)}(t, x)) = \text{prod}(x) - \frac{\text{prod}(x)}{[1 - \text{prod}(t, x)]^2} \\ FD_{(b_k^{(1)}, b_k^{(2)})} (P^{(1)}(t, x) \| P^{(2)}(t, x)) = \frac{\text{prod}(t, x)}{x_k} - \frac{\text{prod}(t, x)}{x_k [1 - \text{prod}(t, x)]^2} \end{cases} \quad (110)$$

with the remainder terms $R_0 (P^{(1)}(t, x) \| P^{(2)}(t, x)) = R_k (P^{(1)}(t, x) \| P^{(2)}(t, x)) = 0, k = 1, 2, \dots, d$. Furthermore, we can get the generalized De Bruijn identities

$$\begin{cases} \frac{\partial}{\partial t} JD (P^{(1)}(t, x), P^{(2)}(t, x)) = -\frac{1}{2} FD_{(b_0^{(1)}, b_0^{(2)})} (P^{(1)}(t, x) \| P^{(2)}(t, x)) \\ \frac{\partial}{\partial x_k} JD (P^{(1)}(t, x), P^{(2)}(t, x)) = -\frac{1}{2} FD_{(b_k^{(1)}, b_k^{(2)})} (P^{(1)}(t, x) \| P^{(2)}(t, x)) \\ k = 1, 2, \dots, d \end{cases} \quad (111)$$

4.2. An Nontrivial Equation

If we let

$$\begin{cases} a_k(u; t, x) = 0 \\ b_k(u; t, x) = \frac{C_1(t, x)C_2^2(t, x)}{4} (C_4(t, x) - u)^2 > 0 \\ k = 0, 1, 2, \dots, d \end{cases} \quad (112)$$

are continuous derivable functions and

$$v(u; t, x) = \frac{1}{C_2(t, x)} \log \left[\frac{C_1(t, x)C_2^2(t, x)}{4C_3(t, x)} (C_4(t, x) - u)^2 \right] \quad (113)$$

is the transformation. Then the Fokker-Planck equations can be rewrite as

$$\frac{\partial p}{\partial t} = \frac{C_1(t, x)}{2} \frac{\partial^2 p}{\partial v^2} + C_2(t, x) \frac{\partial p}{\partial v} + \frac{C_1(t, x)C_2^2(t, x)}{4} p \quad (114)$$

Using Fourier transform, we can get the solution

$$\begin{cases} p(v; t, x) = \frac{e^{\frac{1}{4} \int_0^t C_1(s, x) C_2^2(s, x) ds}}{\sqrt{2\pi \int_0^t C_1(s, x) ds}} e^{-\frac{[v - v_0(x) + \int_0^t C_2(s, x) ds]^2}{2 \int_0^t C_1(s, x) ds}} \\ p(u; t, x) = \frac{e^{\frac{1}{4} \int_0^{x_k} C_1(t, x) C_2^2(t, x) dx_k}}{\sqrt{2\pi \int_0^{x_k} C_1(t, x) dx_k}} e^{-\frac{[v - v_0(x) + \int_0^{x_k} C_2(t, x) dx_k]^2}{2 \int_0^{x_k} C_1(t, x) dx_k}} \\ k = 1, 2, \dots, d \end{cases} \quad (115)$$

where $v_0(x) = \frac{1}{C_2(t,x)} \log \left\{ \frac{C_1(t,x)C_2^2(t,x)}{4C_3(t,x)} [C_4(t,x) - u_0(x)]^2 \right\}$ is the logarithm of initial value. Then we need to suppose that there are functions subject to

$$\begin{cases} \int_0^t C_1(s,x)dt = \int_0^{x_k} C_1(t,x)dx_k \\ \int_0^t C_2(s,x)ds = \int_0^{x_k} C_2(t,x)dx_k \\ \int_0^t C_1(s,x)C_2^2(s,x)ds = \int_0^{x_k} C_1(t,x)C_2^2(t,x)dx_k \\ k = 1, 2, \dots, d \end{cases} \quad (116)$$

and we take $v = v(u; t, x)$ into $p(v; t, x)$, we can get a unified density function formula.

As a example, we let

$$\begin{cases} a_k(u; t, x) \equiv 0 \\ b_k(u; t, x) = b_k(t, x)u^2 > 0 \\ k = 0, 1, 2, \dots, d \end{cases} \quad (117)$$

and $u_0(x) = 1$, the Fokker-Planck equations are

$$\begin{cases} \frac{\partial p^{(3)}}{\partial t} = \frac{b_0(t, x)}{2} u^2 \frac{\partial^2 p^{(3)}}{\partial u^2} + 2b_0(t, x)u \frac{\partial p^{(3)}}{\partial u} + b_0(t, x)p^{(3)} \\ \frac{\partial p^{(3)}}{\partial x_k} = \frac{b_k(t, x)}{2} u^2 \frac{\partial^2 p^{(3)}}{\partial u^2} + 2b_k(t, x)u \frac{\partial p^{(3)}}{\partial u} + b_k(t, x)p^{(3)} \\ k = 1, 2, \dots, d \end{cases} \quad (118)$$

and $v = \log u$ i.e. $u = e^v$, we can obtain

$$\begin{cases} \frac{\partial p^{(3)}}{\partial t} = \frac{b_0(t, x)}{2} \frac{\partial^2 p^{(3)}}{\partial v^2} + \frac{3b_0(t, x)}{2} \frac{\partial p^{(3)}}{\partial v} + b_0(t, x)p^{(3)} \\ \frac{\partial p^{(3)}}{\partial x_k} = \frac{b_k(t, x)}{2} \frac{\partial^2 p^{(3)}}{\partial v^2} + \frac{3b_k(t, x)}{2} \frac{\partial p^{(3)}}{\partial v} + b_k(t, x)p^{(3)} \\ k = 1, 2, \dots, d \end{cases} \quad (119)$$

with the solution

$$\begin{cases} p^{(3)}(v; t, x) = \frac{e^{\int_0^t b_0(s,x)ds}}{\sqrt{2\pi \int_0^t b_0(s,x)ds}} e^{-\frac{[v + \frac{3}{2} \int_0^t b_0(s,x)ds]^2}{2 \int_0^t b_0(s,x)ds}} \\ p^{(3)}(v; t, x) = \frac{e^{\int_0^{x_k} b_k(t,x)dx_k}}{\sqrt{2\pi \int_0^{x_k} b_k(t,x)dx_k}} e^{-\frac{[v + \frac{3}{2} \int_0^{x_k} b_k(t,x)dx_k]^2}{2 \int_0^{x_k} b_k(t,x)dx_k}} \\ k = 1, 2, \dots, d \end{cases} \quad (120)$$

and

$$\begin{cases} p^{(3)}(u; t, x) = \frac{e^{\int_0^t b_0(s,x)ds}}{\sqrt{2\pi \int_0^t b_0(s,x)ds}} e^{-\frac{[\log u + \frac{3}{2} \int_0^t b_0(s,x)ds]^2}{2 \int_0^t b_0(s,x)ds}} \\ p^{(3)}(u; t, x) = \frac{e^{\int_0^{x_k} b_k(t,x)dx_k}}{\sqrt{2\pi \int_0^{x_k} b_k(t,x)dx_k}} e^{-\frac{[\log u + \frac{3}{2} \int_0^{x_k} b_k(t,x)dx_k]^2}{2 \int_0^{x_k} b_k(t,x)dx_k}} \\ k = 1, 2, \dots, d \end{cases} \quad (121)$$

Then we suppose that there is a function $\beta(t, x)$ subject to the total differential

$$d\beta(t, x) = b_0 dt + b_1 dx_1 + \cdots + b_d dx_d \quad (122)$$

and the density function is

$$p^{(3)}(u; t, x) = \frac{e^{\beta(t, x)}}{\sqrt{2\pi\beta(t, x)}} e^{-\frac{[\log u + \frac{3}{2}\beta(t, x)]^2}{2\beta(t, x)}} \quad (123)$$

Similar to the idea of **Section 4.1**, if we consider different $\beta_3(t, x)$ and $\beta_4(t, x)$ in (123), we can get density functions $p^{(3)}(u; t, x)$ and $p^{(4)}(u; t, x)$, then we can get the Jeffreys divergence

$$JD\left(P^{(3)}(t, x), P^{(3)}(s, y)\right) = \frac{\beta_3(t, x) + \beta_3(s, y) + 4}{8\beta_3(t, x)\beta_3(s, y)} [\beta_3(t, x) - \beta_3(s, y)]^2 \quad (124)$$

and generalized Fisher information

$$\begin{cases} FI_1^{(t)}\left(P^{(3)}(t, x)\right) = \frac{\beta_3(t, x) + 2}{4\beta_3^2(t, x)} [b_0^{(3)}(t, x)]^2 \\ FI_1^{(x_k)}\left(P^{(3)}(t, x)\right) = \frac{\beta_3(t, x) + 2}{4\beta_3^2(t, x)} [b_k^{(3)}(t, x)]^2 \end{cases} \quad (125)$$

then the quotients

$$\begin{cases} \frac{JD\left(P^{(3)}(t, x), P^{(3)}(s, y)\right)}{|t - s|^2} = \frac{\beta_3(t, x) + \beta_3(s, y) + 4}{8\beta_3(t, x)\beta_3(s, y)} \left[\frac{\beta_3(t, x) - \beta_3(s, y)}{t - s}\right]^2 \\ \frac{JD\left(P^{(3)}(t, \tilde{x}_k), P^{(3)}(t, \tilde{y}_k)\right)}{|x_k - y_k|^2} = \frac{\beta_3(t, \tilde{x}_k) + \beta_3(t, \tilde{y}_k) + 4}{8\beta_3(t, \tilde{x}_k)\beta_3(t, \tilde{y}_k)} \left[\frac{\beta_3(t, \tilde{x}_k) - \beta_3(t, \tilde{y}_k)}{x_k - y_k}\right]^2 \end{cases} \quad (126)$$

and we can easily get the relation

$$\begin{cases} \frac{JD\left(P^{(3)}(t, x), P^{(3)}(s, y)\right)}{|t - s|^2 FI_1^{(t)}\left(P^{(3)}(t, x)\right)} = \frac{\beta_3(t, x)}{\beta_3(s, x)} \frac{\beta_3(t, x) + \beta_3(s, y) + 4}{2[\beta_3(t, x) + 2]} \left[\frac{\beta_3(t, x) - \beta_3(s, y)}{b_0^{(3)}(t, x)(t - s)}\right]^2 \\ \frac{JD\left(P^{(3)}(t, \tilde{x}_k), P^{(3)}(t, \tilde{y}_k)\right)}{|x_k - y_k|^2 FI_1^{(x_k)}\left(P^{(3)}(t, \tilde{x}_k)\right)} = \frac{\beta_3(t, \tilde{x}_k)}{\beta_3(t, \tilde{y}_k)} \frac{\beta_3(t, \tilde{x}_k) + \beta_3(t, \tilde{y}_k) + 4}{2[\beta_3(t, \tilde{x}_k) + 2]} \left[\frac{\beta_3(t, \tilde{x}_k) - \beta_3(t, \tilde{y}_k)}{b_k^{(3)}(t, x)(x_k - y_k)}\right]^2 \end{cases} \quad (127)$$

for $k = 1, 2, \dots, d$. Without loss of generality, the result (127) also corroborates **Theorem 4**.

Furthermore, if we consider different $\beta_3(t, x)$ and $\beta_4(t, x)$ in (123), we can get 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

$$\begin{aligned} & FD_{(b_k^{(3)}, b_k^{(4)})}\left(P^{(3)}(t, x) \| P^{(4)}(t, x)\right) \\ &= \frac{\beta_3(t, x) - \beta_4(t, x)}{4\beta_3^2(t, x)\beta_4^2(s, y)} \\ &\times \left\{ b_k^{(4)}(t, x)\beta_3(t, x) [\beta_3^2(t, x) - 2\beta_4^2(s, y)] - b_k^{(3)}(t, x)\beta_4(t, x) [\beta_4^2(s, y) - 2\beta_3^2(t, x)] \right\} \\ &\times [4\beta_3(t, x) + 4\beta_4(t, x) - 3\beta_3(t, x)\beta_4(t, x)] [b_k^{(4)}(t, x)\beta_3(t, x) - b_k^{(3)}(t, x)\beta_4(t, x)] \end{aligned} \quad (128)$$

with the remainder terms $R_k \left(P^{(3)}(t, x) \| P^{(4)}(t, x) \right) = 2 \left[b_k^{(3)}(t, x) - b_k^{(4)}(t, x) \right], k = 0, 1, 2, \dots, d$. Then the generalized De Bruijn identities are

$$\begin{cases} \frac{\partial}{\partial t} JD \left(P^{(3)}(t, x), P^{(4)}(t, x) \right) = -\frac{1}{2} FD_{(b_0^{(3)}, b_0^{(4)})} \left(P^{(3)}(t, x) \| P^{(4)}(t, x) \right) - 2 \left[b_0^{(3)}(t, x) - b_0^{(4)}(t, x) \right] \\ \frac{\partial}{\partial x_k} JD \left(P^{(3)}(t, x), P^{(4)}(t, x) \right) = -\frac{1}{2} FD_{(b_k^{(3)}, b_k^{(4)})} \left(P^{(3)}(t, x) \| P^{(4)}(t, x) \right) - 2 \left[b_k^{(3)}(t, x) - b_k^{(4)}(t, x) \right] \\ k = 1, 2, \dots, d \end{cases} \quad (129)$$

4.3. An Interesting Equation

If we let

$$\begin{cases} a_k(u) = -\frac{3D_1^2 D_3}{8} (u - D_6) \\ b_k(u) = D_1^2 D_5^2 \left[1 - \frac{D_3}{4D_5^2} (u - D_6)^2 \right] \\ k = 0, 1, 2, \dots, d \end{cases} \quad (130)$$

are continuous derivable functions and

$$u = \frac{2D_5}{\sqrt{D_3}} \sin \left(\frac{\sqrt{D_3}}{2} v + D_4 \right) + D_6 \quad (131)$$

where D_1 to D_6 are nonnegative functions independent of u and we omit (t, x) .

Thus, we can rewrite the Fokker-Planck equations as

$$\begin{cases} \frac{\partial p}{\partial t} = \frac{D_1^2}{2} \frac{\partial^2 p}{\partial v^2} + \frac{D_1^2 D_3}{8} p \\ \frac{\partial p}{\partial x_k} = \frac{D_1^2}{2} \frac{\partial^2 p}{\partial v^2} + \frac{D_1^2 D_3}{8} p \\ k = 1, 2, \dots, d \end{cases} \quad (132)$$

with the solution

$$p(v; t, x) = \frac{e^{\frac{1}{8} \int_0^t D_1^2 D_3 ds} e^{-\frac{(v-v_0)^2}{2 \int_0^t D_1^2 ds}}}{\sqrt{2\pi \int_0^t D_1^2 ds}} \quad (133)$$

Recall the **Lemma 2**, we obtain

$$\begin{aligned} & \int_{I_u} p(u; t, x) du \\ &= \frac{D_5 e^{\frac{1}{8} \int_0^t D_1^2 D_3 ds}}{\sqrt{2\pi \int_0^t D_1^2 ds}} \int_{\mathbb{R}} \cos \left(\frac{\sqrt{D_3}}{2} \tilde{u} + D_4 \right) e^{-\frac{(\tilde{u}-\tilde{u}_0)^2}{2 \int_0^t D_1^2 ds}} d\tilde{u} \\ &= 1 \end{aligned} \quad (134)$$

and use the result

$$\frac{1}{\sqrt{2\pi H}} \int_{\mathbb{R}} \cos(\omega x + \varphi) e^{-\frac{(x-G)^2}{2H}} dx = e^{-\frac{\omega^2 H}{2}} \cos(\omega G + \varphi) \quad (135)$$

the integral (134) can be write as

$$D_5 e^{\frac{1}{8} \int_0^t D_1^2 D_3 ds - \frac{D_3}{8} \int_0^t D_1^2 ds} \cos \left(\frac{\sqrt{D_3}}{2} \tilde{u}_0 + D_4 \right) = 1 \quad (136)$$

where $\tilde{u}_0 = \frac{2}{\sqrt{D_3}} \arcsin \left[\frac{\sqrt{D_3}}{2D_5} (u_0 - D_6) \right]$.

Here we give a simple example. Let $u \in [-1, 1]$,

$$\begin{cases} a_k(u; t, x) = -\frac{3}{2} b_k(t, x) u \\ b_k(u; t, x) = b_k(t, x) (1 - u^2) \\ k = 0, 1, 2, \dots, d \end{cases} \quad (137)$$

where $b_k(t, x) > 0$ and we can obtain the Fokker-Planck equations formed as

$$\begin{cases} \frac{\partial p}{\partial t} = \frac{b_0(t, x)}{2} \frac{\partial^2}{\partial u^2} \left[(1 - u^2) p \right] - \frac{3b_0(t, x)}{2} \frac{\partial}{\partial u} (up) \\ \frac{\partial p}{\partial x_k} = \frac{b_k(t, x)}{2} \frac{\partial^2}{\partial u^2} \left[(1 - u^2) p \right] - \frac{3b_k(t, x)}{2} \frac{\partial}{\partial u} (up) \\ k = 1, 2, \dots, d \end{cases} \quad (138)$$

with initial value $u_0(x) = 0$.

Using the formula $u = \sin v$, we can rewrite the Fokker-Planck equations

$$\begin{cases} \frac{\partial p}{\partial t} = \frac{b_0(t, x)}{2} \frac{\partial^2 p}{\partial v^2} + \frac{b_0(t, x)}{2} p \\ \frac{\partial p}{\partial x_k} = \frac{b_k(t, x)}{2} \frac{\partial^2 p}{\partial v^2} + \frac{b_k(t, x)}{2} p \\ k = 1, 2, \dots, d \end{cases} \quad (139)$$

with the solution

$$\begin{cases} p(v; t, x) = \frac{e^{\frac{1}{2} \int_0^t b_0(s, x) ds}}{\sqrt{2\pi \int_0^t b_0(s, x) ds}} e^{-\frac{v^2}{2 \int_0^t b_0(s, x) ds}} \\ p(v; t, x) = \frac{e^{\frac{1}{2} \int_0^{x_k} b_k(t, x) dx_k}}{\sqrt{2\pi \int_0^{x_k} b_k(t, x) dx_k}} e^{-\frac{v^2}{2 \int_0^{x_k} b_k(t, x) dx_k}} \end{cases} \quad (140)$$

and the variable substitution is $u = \sin v$.

Similarly, we suppose that there is a function $\beta(t, x)$ subject to the total differential

$$d\beta(t, x) = b_0 dt + b_1 dx_1 + \dots + b_d dx_d \quad (141)$$

then the density is

$$\begin{cases} p(v; t, x) = \frac{e^{\frac{1}{2} \beta(t, x)}}{\sqrt{2\pi \beta(t, x)}} e^{-\frac{v^2}{2\beta(t, x)}} \\ u = \sin v \end{cases} \quad (142)$$

From density function (142), if we consider different $\beta_5(t, x)$ and $\beta_6(t, x)$ in (142), we can get density functions $p^{(5)}(u; t, x)$ and $p^{(6)}(u; t, x)$, then we can get the Jeffreys divergence and generalized Fisher information

$$JD \left(P^{(5)}(t, x), P^{(5)}(s, y) \right) = \frac{1 - \beta_5(t, x) - \beta_5(s, y)}{2\beta_5(t, x)\beta_5(s, y)} [\beta_5(t, x) - \beta_5(s, y)]^2 \quad (143)$$

and

$$\begin{cases} FI_1^{(t)}(P^{(5)}(t, x)) = \frac{1 - 2\beta_5(t, x)}{2\beta_5^2(t, x)} [b_0^{(5)}(t, x)]^2 \\ FI_1^{(x_k)}(P^{(5)}(t, x)) = \frac{1 - 2\beta_5(t, x)}{2\beta_5^2(t, x)} [b_k^{(5)}(t, x)]^2 \end{cases} \quad (144)$$

then the quotients

$$\begin{cases} \frac{JD(P^{(5)}(t, x), P^{(5)}(s, x))}{|t - s|^2} = \frac{1 - \beta_5(t, x) - \beta_5(s, x)}{2\beta_5(t, x)\beta_5(s, x)} \left[\frac{\beta_5(t, x) - \beta_5(s, x)}{t - s} \right]^2 \\ \frac{JD(P^{(5)}(t, \tilde{x}_k), P^{(5)}(t, \tilde{y}_k))}{|\tilde{x}_k - \tilde{y}_k|^2} = \frac{1 - \beta_5(t, \tilde{x}_k) - \beta_5(t, \tilde{y}_k)}{2\beta_5(t, \tilde{x}_k)\beta_5(t, \tilde{y}_k)} \left[\frac{\beta_5(t, \tilde{x}_k) - \beta_5(t, \tilde{y}_k)}{\tilde{x}_k - \tilde{y}_k} \right]^2 \end{cases} \quad (145)$$

Obviously, we can easily get

$$\begin{cases} \frac{JD(P^{(5)}(t, x), P^{(5)}(s, x))}{|t - s|^2 FI_1^{(t)}(P^{(5)}(t, x))} = \frac{1 - \beta_5(t, x) - \beta_5(s, x)}{1 - 2\beta_5(t, x)} \frac{\beta_5(t, x)}{\beta_5(s, x)} \left[\frac{\beta_5(t, x) - \beta_5(s, x)}{b_0^{(5)}(t, x)(t - s)} \right]^2 \\ \frac{JD(P^{(5)}(t, \tilde{x}_k), P^{(5)}(t, \tilde{y}_k))}{|\tilde{x}_k - \tilde{y}_k|^2 FI_1^{(x_k)}(P^{(5)}(t, x))} = \frac{1 - \beta_5(t, \tilde{x}_k) - \beta_5(t, \tilde{y}_k)}{1 - 2\beta_5(t, \tilde{x}_k)} \frac{\beta_5(t, \tilde{x}_k)}{\beta_5(t, \tilde{y}_k)} \left[\frac{\beta_5(t, \tilde{x}_k) - \beta_5(t, \tilde{y}_k)}{b_k^{(5)}(t, x)(\tilde{x}_k - \tilde{y}_k)} \right]^2 \end{cases} \quad (146)$$

for $k = 1, 2, \dots, d$. Without loss of generality, the result (146) corroborates **Theorem 4**.

Furthermore, if we consider different $\beta_5(t, x)$ and $\beta_6(t, x)$ in (142), we can get density functions $p^{(5)}(u; t, x)$ and $p^{(6)}(u; t, x)$, then the generalized Fisher divergence at the same space-time points is

$$\begin{aligned} & FD_{(b_k^{(5)}, b_k^{(6)})} (P^{(5)}(t, x) \| P^{(6)}(t, x)) \\ &= \frac{[\beta_5(t, x) - \beta_6(t, x)] [b_k^{(6)} \beta_5(t, x) - b_k^{(5)} \beta_6(t, x)]}{2\beta_5^2(t, x)\beta_6^2(t, x)} [\beta_5(t, x) - \beta_5^2(t, x) + \beta_6(t, x) - \beta_6^2(t, x)] \end{aligned} \quad (147)$$

with the remainder terms $R_k(P^{(5)}(t, x) \| P^{(6)}(t, x)) = b_k^{(5)}(t, x) - b_k^{(6)}(t, x)$, $k = 0, 1, 2, \dots, d$. Then the generalized De Bruijn identities are

$$\begin{cases} \frac{\partial}{\partial t} JD(P^{(5)}(t, x), P^{(6)}(t, x)) = -\frac{1}{2} FD_{(b_0^{(5)}, b_0^{(6)})} (P^{(5)}(t, x) \| P^{(6)}(t, x)) - [b_0^{(5)}(t, x) - b_0^{(6)}(t, x)] \\ \frac{\partial}{\partial x_k} JD(P^{(5)}(t, x), P^{(6)}(t, x)) = -\frac{1}{2} FD_{(b_k^{(5)}, b_k^{(6)})} (P^{(5)}(t, x) \| P^{(6)}(t, x)) - [b_k^{(5)}(t, x) - b_k^{(6)}(t, x)] \\ k = 1, 2, \dots, d \end{cases} \quad (148)$$

5. Conclusions

In this paper, we generalize the classical definitions of entropy, divergence and Fisher information, obtain these measures on space-time random field. In addition to that, we also get the Fokker-Planck equations (55) for space-time random field and obtain the density functions. Moreover, we obtain the Jeffreys divergence of a space-time random field at different space-time positions, and we get the approximation of the ratio of Jeffreys divergence to the square of space-time coordinate difference to the generalized Fisher information(68). Further, we use the Jeffreys divergence on two space-time random fields from same type but different parameters Fokker-Planck equations, to obtain generalized De Bruijn identities (77), and get the relation between Jeffreys divergence of space-time random field and generalized Fisher divergence. Finally, we give three examples of Fokker-Planck equations, with their solutions, 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.

Acknowledgments: The author would like to thank Prof. Pingyi Fan for providing relevant references and helpful discussions on topics related to this work.

Conflicts of Interest: The authors 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*, Berlin: Springer Heidelberg, 1984.
2. Stam, A. J. Some inequalities satisfied by the quantities of information of Fisher and Shannon. *Inf. Control* 1959, 2, 101–112.
3. Barron, A. R. Entropy and the central limit theorem. *Ann. Probab.* 1986, 14, 336–342.
4. Johnson, O. *Information theory and the central limit theorem*, London, U.K.: Imperial college Press, 2004.
5. Guo, D. Relative entropy and score function: New information estimation relationships through arbitrary additive perturbation, in Proc. IEEE Int. Symp. Inf. Theory, Seoul, South Korea, Jun./Jul. 2009, 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.
7. Kharazmi, O.; Balakrishnan, N. Cumulative residual and relative cumulative residual Fisher information and their properties. *IEEE Trans. Inform. Theory* 2021, 67, 6306–6312.
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.
12. Yaglom, A. M. *Correlation theory of stationary and related Random functions. Volume I: basic results*, New York: Springer-Verlag, 1987.
13. Yaglom, A. M. *Correlation theory of stationary and related random functions. Volume II: supplementary notes and references*, Berlin: Springer-Velag, 1987.
14. Bowditch, A.; Sun, R. The two-dimensional continuum random field Ising model. *Ann. Probab.* 2022, 50, 419–454.
15. Bailleul, I.; Catellier, R.; Delarue, F. Propagation of chaos for mean field rough differential equations. *Ann. Probab.* 2021, 49, 944–996.
16. Wu, L.; Samorodnitsky, G. Regularly varying random fields. *Stoch. Process Their. Appl.* 2020, 130, 4470–4492.
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.
18. Ye, Z. On entropy and ε -entropy of random fields, Ph.D.dissertation. Cornell University, 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.
20. Ye, Z.; Berger, T. *Information Measures for Discrete Random Fields*, Beijing/New York: Science Press, 1998.
21. Ye, Z.; Yang, W. *Random Field: Network Information Theory and Game Theory (in Chinese)*, Beijing: Science Press, 2023.

22. Ma, C. Stationary random fields in space and time with rational spectral densities. *IEEE Trans. Inform. Theory* 2007, 53, 1019–1029.
23. Hairer, M. A theory of regularity structures. *Invent. Math.* 2014, 198, 269–504.
24. Hairer, M. Solving the KPZ equation. *Ann. Math.* 2013, 178, 559–664.
25. Kremp, H.; Perkowski, N. Multidimensional SDE with distributional drift and Lévy noise. *Bernoulli* 2022, 28, 1757–1783.
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, 3, 3054–3090.
27. Song, Z.; Zhang, J. A note for estimation about average differential entropy of continuous bounded space-time random field. *Chinese J. Electron* 2022, 31, 793–803.
28. Kramers, H. A. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* 1940, 7, 284–304.
29. Moyal, J. E. Stochastic processes and statistical physics. *J R Stat Soc Series B Stat Methodol* 1949, 11, 150–210.
30. Shannon, C. E. A mathematical theory of communication. *Bell System Technical Journal* 1948, 27, 379–423, 623–656.
31. Neeser, F. D.; Massey, J. L. Proper complex random processes with applications to information theory. *IEEE Trans. Inform. Theory* 1991, 39, 1293–1302.
32. Ihara, S. *Information theory-for continuous systems*, Singapore: World Scientific, 1993.
33. Gray, R. M. *Entropy and information theory*, Boston: Springer, 2011.
34. Bach, F. Information Theory With Kernel Methods. *IEEE Trans. Inform. Theory* 2023, 69, 752–775.
35. Kullback, S.; Leibler, R. A. On information and sufficiency, *Ann. Math. Stat.* 1951, 22, 79–86.
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; p. 31.
38. Rényi, A. On measures of entropy and information. Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, 1961, 1, 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.
40. Liu, S.; She, R.; Zhu, Z.; Fan, P. Storage Space Allocation Strategy for Digital Data with Message Importance. *Entropy* 2020, 22, 591.
41. She, R.; Liu, S.; Fan, P. Attention to the Variation of Probabilistic Events: Information Processing with Message Importance Measure. *Entropy* 2019, 21, 439.
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.
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.
44. Nielsen, F. An Elementary Introduction to Information Geometry. *Entropy* 2020, 22, 1100.
45. Nielsen, F. On the Jensen–Shannon Symmetrization of Distances Relying on Abstract Means. *Entropy* 2019, 21, 485.
46. Nielsen, F.; Nock, R. Generalizing skew Jensen divergences and Bregman divergences with comparative convexity. *IEEE Signal Process. Lett.* 2017, 24, 1123–1127.
47. Furuichi, S.; Minculete, N. Refined Young Inequality and Its Application to Divergences. *Entropy* 2021, 23, 514.
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.
49. Reverter, F.; Oller, J.M. Computing the Rao distance for Gamma distributions. *J. Comput. Appl. Math.* 2003, 157, 155–167.
50. Pawula, R. F. Generalizations and extensions of the Fokker-Planck-Kolmogorov equations. *IEEE Trans. Inform. Theory* 1967, 13, 33–41.
51. Pawula, R. F. Approximation of the linear Boltzmann equation by the Fokker-Planck equation. *Phys. Rev.* 1967, 162, 186–188.

52. Khoshnevisan, D.; Shi, Z. Brownian Sheet and Capacity, *Ann. Probab.* 1999, 27, 3, 1135–1159.
53. Daniel, R.; Marc. Y. *Continuous Martingales and Brownian Motion 2nd*. New York: Springer-Verlag. 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.