

Article

Not peer-reviewed version

Information Geometry Description of Inferential Scattering

[Marco Favretti](#)*

Posted Date: 14 April 2026

doi: 10.20944/preprints202604.0922.v1

Keywords: Riemannian manifold; dually flat manifold; exponential family; linear connection; scattering theory



Preprints.org is a free multidisciplinary 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 open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

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.

Article

Information Geometry Description of Inferential Scattering

Marco Favretti

Dipartimento di Matematica "Tullio Levi-Civita", Università degli Studi di Padova, Padova, Italy; favretti@math.unipd.it

Abstract

We investigate the geometrical structure underlying the notion of Inferential Scattering, which was formulated by E. T. Jaynes in the 1980s using the language of equilibrium statistical mechanics. We show that inferential scattering can be naturally defined on a dually flat Riemannian manifold equipped with dual coordinate systems, a differential-geometric structure that occupies a central place in information geometry. We find that the evolution of the system on the dually flat manifold can be expressed as the horizontal lift of an integrable connection. We stress that the notion of inferential scattering has a wide range of applications, being a form of inference and therefore applicable to any statistical system with insufficient information.

Keywords: Riemannian manifold; dually flat manifold; exponential family; linear connection; scattering theory

1. Introduction

The notion of Inferential Scattering was exposed by E.T. Jaynes in a couple of papers [1,2] whose content has remained little known in the scientific literature with respect to more celebrated works of this author. Jaynes' original exposition refers to equilibrium statistical mechanics. It is well known that the probabilistic formalism of equilibrium statistical mechanics is described by an exponential family manifold which is an instance of a dually flat manifold [3]; the main aim of this work is to rephrase the content of the notion of inferential scattering in this enlarged setting of dually flat manifolds to better appreciate its geometric structure and possibly expand its range of applications. Dually flat manifolds are a generalization of Riemannian manifolds and they occupy a central place in Information Geometry. They have found multiple applications in the description of neural networks [4], machine learning (especially Restricted Boltzmann Machines) [5] and the description of hierarchical systems [6].

In a dually flat manifold M there are two local coordinate systems $\theta = \theta(p)$ and $\eta = \eta(p)$ which describe the same point $p \in M$. The transition maps $\eta \circ \theta^{-1}$ and $\theta \circ \eta^{-1}$ are related by a Legendre transform so that $\eta(\theta) = \nabla\psi(\theta)$ with inverse $\theta(\eta) = \nabla\varphi(\eta)$ where ∇ denotes the gradient and φ, ψ are called potentials. In the case of an exponential manifold, there is a natural statistical mechanic interpretation of θ as control variables and η as effect variables, see Section 4. The tangents of the above introduced gradient maps give the linear variation $d\eta$ in the effect variables due to a variation $d\theta$ in control variables and vice-versa.

The phenomenon of inferential scattering arises when particular kinds of variations $d\eta$ and $d\theta$ are considered. In this work it is shown that its underlying geometric structure can be understood in terms of horizontal lift with respect to a couple of integrable connections (dually orthogonal foliations) which are characteristics of the dually flat manifold structure. The notion of inferential scattering comes into play because the variation $d\eta$ in effect variables can be written as an infinite sum of contributions. According to Jaynes [1] this perturbative expansion can be interpreted using the multiple scattering phenomenon of a wave emanating from a point source A in presence of an obstacle C : the intensity of the wave at a point B in space is the sum of the contribution of the wave arriving from A plus the contribution of the wave reflected by C plus the one of the wave reflected from C back in A and then

in B and so on with the higher order multiple reflections (Born series, see [10]). The main interest of this analysis of Jaynes' result is to show that the scattering phenomenon is present also in this abstract setting of a dually flat manifold, where it establishes a relation between the dual coordinate systems θ and η .

We conclude with the plan of the paper: In Section 2 we give a cursory view of the differential geometry notion of dually flat manifold and mutually dual foliations. In Section 3 we describe the paradigmatic case of the exponential family manifold and in Section 4 we investigate the geometric structure of the notion of direct and inverse control and in Section 5 the one of inferential scattering.

2. Dually Flat Manifolds

We briefly recall the notions of differential geometry useful for introducing the definition of a dually flat manifold, a central object in information geometry. We keep the exposition to a minimum, without proofs, essentially following [3]. Other important introductory references are [5,7,9].

Let (M, g) be a n -dimensional smooth manifold with g a Riemannian metric. In a local chart ζ the coordinates of a point $p \in M$ are $\zeta(p) = (x_1, \dots, x_n)$ and the canonical basis is $\partial_i = \frac{\partial}{\partial x_i}$. We equip M with a linear connection ∇ whose expression in the local chart is given by the Christoffel symbols Γ_{ij}^k as

$$\nabla_{\partial_i} \partial_j = \Gamma_{ij}^k \partial_k, \quad \Gamma_{ij,k} = g(\nabla_{\partial_i} \partial_j, \partial_k) = g_{lk} \Gamma_{ij}^l.$$

Throughout the paper we adopt the convention that repeated indices are summed. The torsion τ of the connection ∇ is defined as

$$\tau(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y]$$

where $X, Y \in \mathfrak{X}(M)$ are vector fields on M and $[X, Y]$ is the Lie bracket. The connection ∇ is said to be compatible with the metric g or simply metric if

$$X \cdot g(X, Y) = g(\nabla_X Y, Z) + g(X, \nabla_Y Z) \quad \forall X, Y, Z \in \mathfrak{X}(M). \quad (1)$$

If the connection is metric, the parallel transport defined by ∇ is an isometry, preserving the length and the angle between vectors. A celebrated theorem in differential geometry states that in a Riemannian manifold there exists a unique metric connection which is torsion-free $\tau = 0$, the so-called Levi-Civita connection $\hat{\nabla}$. For later use we introduce the curvature defined by the connection ∇

$$R(X, Y, Z) = \nabla_X(\nabla_Y Z) - \nabla_Y(\nabla_X Z) - \nabla_{[X, Y]} Z.$$

The connection ∇ is flat if the associated curvature is identically zero. This is a coordinate-free condition and the Christoffel symbols are zero in every coordinate chart. Moreover, there exists a local affine coordinate system with the property that every canonical vector $\partial/\partial x_i$ of the basis can be parallel transported along the coordinate line x_i so each coordinate line is a geodesic.

The above notion (1) of compatibility between the metric g and a connection ∇ is generalized in Information Geometry using the following

Definition 1. *The connections ∇ and ∇^* are called conjugate or dual with respect to the metric in (M, g) if*

$$X \cdot g(X, Y) = g(\nabla_X Y, Z) + g(X, \nabla_Y^* Z) \quad \forall X, Y, Z \in \mathfrak{X}(M).$$

One can show [3] that the connection ∇^* dual to ∇ is unique. Moreover, in [3] it is proved that if ∇, ∇^* are dual and both torsion-free, then $R = 0$ if and only if $R^* = 0$. This motivates the following

Definition 2. *A Riemannian manifold (M, g, ∇, ∇^*) equipped with two connections which are dual and torsion-free is called dually flat when $R = R^* = 0$, i.e. both connections are flat.*

As a consequence, there exists on M two affine local coordinate systems θ and η with the property that the vectors $\partial^i = \frac{\partial}{\partial \eta^i}$ of the basis are ∇ -parallel transported along the coordinate lines η^i and the vectors $\partial_i = \frac{\partial}{\partial \theta^i}$ are ∇^* -parallel transported along the coordinate lines θ_i . The coordinates θ and η are called dual or bi-orthogonal because it holds that $g(\partial_i, \partial^j) = \delta_i^j$ where δ is the Kronecker symbol. Note that the two basis are orthogonal each other but the vectors of each base need not to be orthogonal. More in detail, if we consider the representation of the metric g in the two basis

$$g^{ij} = g(\partial^i, \partial^j), \quad g_{ij} = g(\partial_i, \partial_j)$$

it holds that the two matrices are in the relation $g^{ij}g_{jk} = \delta_k^i$. Finally, let $\theta = \theta(\eta)$ and $\eta = \eta(\theta)$ be the transition maps connecting the representation of a point $p \in M$ in the two coordinate charts. Note that they are one the inverse of the other. Using the orthogonality condition $g^{ij}g_{jk} = \delta_k^i$ one can prove that their tangent maps coincide with the representation of the metric g in the two coordinate basis, that is

$$g^{ij}(\eta) = \frac{\partial \theta_i(\eta)}{\partial \eta^j}, \quad g_{ij}(\theta) = \frac{\partial \eta^i(\theta)}{\partial \theta_j}. \quad (2)$$

We now show that

Proposition 1. *The change of coordinate $\eta = \eta(\theta)$ and $\theta = \theta(\eta)$ can be represented using the so-called potential scalar functions $\varphi(\eta)$ and $\psi(\theta)$ in the sense that*

$$\theta_i(\eta) = \frac{\partial \varphi(\eta)}{\partial \eta^i}, \quad \eta^i(\theta) = \frac{\partial \psi(\theta)}{\partial \theta^i}, \quad i = 1, \dots, n. \quad (3)$$

Proof. We prove the first relation of (3), the proof for the second follows the same pattern. We need to prove that the one form $\omega = \theta_i(\eta)d\eta^i$ is exact, that is there exists a scalar function $\varphi(\eta)$ such that $d\varphi = \omega$. A necessary (and sufficient condition if the manifold is simply connected) is that $d\omega = 0$. Then we need to check that

$$d\omega = d(\theta_i(\eta)d\eta^i) = \frac{\partial \theta_i}{\partial \eta^j} d\eta^j \wedge d\eta^i = \left(\frac{\partial \theta_i}{\partial \eta^j} - \frac{\partial \theta_j}{\partial \eta^i} \right) d\eta^j \otimes d\eta^i = 0$$

that is

$$\frac{\partial \theta_i}{\partial \eta^j} - \frac{\partial \theta_j}{\partial \eta^i} = 0 \quad \forall i, j = 1, \dots, n$$

Now, using (2), the above relation becomes $g^{ij} = g^{ji}$ which is true. \square

Note that the relations (2) and (3) together imply that the potentials $\varphi(\eta)$ and $\psi(\theta)$ are convex function. Indeed we have

$$\frac{\partial^2 \varphi(\eta)}{\partial \eta^j \partial \eta^i} = \frac{\partial \theta_i}{\partial \eta^j} = g^{ij} \in \text{Sym}^+(n)$$

i.e. that the Hessian matrix of φ is positive definite. Since the metric g is the Hessian of a function, the dually flat manifold is an example of an Hessian manifold. We do not pursue this line of investigation, see [9] for further details.

Note that the existence of a dual coordinate system induces stringent conditions on the metric g : indeed by deriving (2) and using Schwartz condition on mixed second order derivative we have that

$$\frac{\partial g^{ij}}{\partial \eta^k} = \frac{\partial g^{ik}}{\partial \eta^j} \quad \text{and} \quad \frac{\partial g_{ij}}{\partial \theta^k} = \frac{\partial g_{ik}}{\partial \theta^j}.$$

Another remarkable consequence is the following, which is a classical result of Legendre transform theory: since the change of coordinate maps are gradient maps, their inverse are also gradient maps and they are related by the Legendre condition

$$\varphi(\eta) = \theta(\eta) \cdot \eta - \psi(\theta(\eta)).$$

Indeed, using the one-form notation, we have

$$d\varphi(\eta) = d\theta \cdot \eta + \theta \cdot d\eta - d\psi = d\theta \cdot \eta + \theta \cdot d\eta - \eta \cdot d\theta = \theta \cdot d\eta.$$

2.1. Mutually Dual Foliations of a Dually Flat Space

A remarkable fact of dually flat manifolds is that they can admit a couple of orthogonal foliations. Let us consider an n -dimensional dually flat manifold (M, g, ∇, ∇^*) and let $\eta(p) = \eta, \theta(p) = \theta, p \in M$, be a couple of bi-orthogonal affine local coordinate systems. Given a k with $1 < k < n$ we partition the coordinates in two subsets

$$\eta = (\eta^A, \eta^B) = (\eta^1, \dots, \eta^k, \eta^{k+1}, \dots, \eta^n) \quad \text{and} \quad \theta = (\theta_A, \theta_B) = (\theta_1, \dots, \theta_k, \theta_{k+1}, \dots, \theta_n)$$

and we consider the coordinate projections $\pi_A(\eta) = \eta^A$ and $\pi_B(\theta) = \theta_B$. For every $\bar{\eta}^A$ and $\bar{\theta}_B$ we define the submanifolds of M

$$E(\bar{\eta}^A) = \{p \in M : \pi_A \circ \eta(p) = \bar{\eta}^A\} \quad Q(\bar{\theta}_B) = \{p \in M : \pi_B \circ \theta(p) = \bar{\theta}_B\} \quad (4)$$

In [3] is proved that since M is dually flat and the subsets $E(\bar{\eta}^A)$ and $Q(\bar{\theta}_B)$ are defined as affine subspaces of an affine coordinate system, they are totally geodesic submanifolds with respect to the ∇ and ∇^* connection respectively, that is it holds that a ∇ -geodesic (respectively ∇^* -geodesic) which is tangent to the submanifold at a point remains tangent to the submanifold. Moreover since their tangent spaces can be described as

$$T_p E(\bar{\eta}^A) = \langle \partial^{k+1}, \dots, \partial^n \rangle, \quad T_p Q(\bar{\theta}_B) = \langle \partial_1, \dots, \partial_k \rangle$$

the two submanifolds are mutually orthogonal by the condition $g(\partial_i, \partial^j) = \delta_i^j$ and complementary in the sense that their direct sum of subspaces gives the tangent space $T_p M$. See Figure 1 below. If we introduce the so called mixed coordinate system

$$m(p) = (\eta^A(p), \theta_B(p)) = (\pi_A(\eta(p)), \pi_B(\theta(p)))$$

the metric g admits the block decomposition

$$g(m(p)) = g^{ab} d\eta^a \otimes d\eta^b + g_{cd} d\theta_c \otimes d\theta_d, \quad a, b \in \{1, \dots, k\}, \quad c, d \in \{k+1, \dots, n\}.$$

So the complementary projections π_A and π_B define a couple of mutually dual foliation of M whose leaves are orthogonal submanifolds of constant dimension (respectively $\dim E(\bar{\eta}^A) = n - k$, and $\dim Q(\bar{\theta}_B) = k$) and totally geodesic with respect to the ∇ and ∇^* connections. This partition of the space is called k -cut of a dual coordinate system in [5]. The notion of mutually dual foliations has important applications on Information Geometry due to the fact that a version of the Pithagorean Theorem applies (see again [3]).

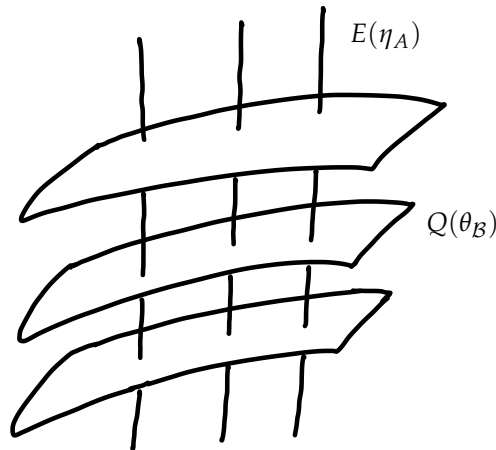


Figure 1. Cartoon picture of a mutually dual foliation in a dually flat space.

3. Exponential Family Manifold

We describe an example of dually flat manifold which occupies a central place in information geometry. We start with the definition of statistical model. Let $(X, \mathcal{B}, d\mu)$ be a measure space, where the set X can be numerable or not, \mathcal{B} is a σ -algebra over X and $d\mu$ is a measure on X . The infinite dimensional space of all the positive probability densities over X is

$$\mathcal{P}(X) = \{p(x) \geq 0 \text{ a.e. on } X, \text{ and } \int_X p(x)d\mu = 1\}.$$

Given an open convex set $U \subset \mathbb{R}^n$ with coordinates $\theta = (\theta_1, \dots, \theta_n)$ we consider the subset of $\mathcal{P}(X)$ of probability densities parameterized by the parameters $\theta \in U$

$$S = \{p_\theta = p(x, \theta), \quad \theta \in U\} \subset \mathcal{P}(X).$$

The set S has the structure of a n -dimensional sub-manifold of $\mathcal{P}(X)$ if: 1) the map $\theta \mapsto p_\theta$ is injective and 2) the n base functions $p_i = \partial p(x, \theta) / \partial \theta_i$ are linearly independent as functions over X . In this case we call S a n -dimensional statistical model. Usually, linear independence condition is checked using the equivalent score function basis functions $l_i = \partial l / \partial \theta_i$, where $l = \ln p$ is the likelihood of p .

An exponential family is the n -dimensional statistical manifold

$$\mathcal{E} = \{p(x, \theta) = e^{\theta \cdot h(x) + k(x) - \psi(\theta)}, \quad \theta \in U\} \subset \mathcal{P}(X) \quad (5)$$

where $h : X \rightarrow \mathbb{R}^n$ is a map of maximum rank n whose components $h_i, i = 1, \dots, n$ are called feature functions or observables, k is a positive function such that $e^{k(x)}d\mu$ is a probability density over X and the function $\psi(\theta)$

$$\psi(\theta) = \ln \int_X e^{\theta \cdot h(x) + k(x)} d\mu$$

is the cumulant or moment generating function, called free energy in statistical mechanics. A direct computation shows that the averages of h_i can be written as

$$\eta_i = \mathbb{E}_{p_\theta}(h_i) = \frac{\partial \psi(\theta)}{\partial \theta_i}, \quad i = 1, \dots, n \quad (6)$$

and that the covariances are

$$\text{cov}(h_i, h_j) = \mathbb{E}_{p_\theta}(h_i h_j) - \mathbb{E}_{p_\theta}(h_i) \mathbb{E}_{p_\theta}(h_j) = \frac{\partial^2 \psi(\theta)}{\partial \theta_i \partial \theta_j}.$$

Using the score vectors $l_i = h_i - \partial\psi/\partial\theta_i$ associated to the exponential model one can show that the Hessian of the cumulant function coincides with the Fisher Riemannian metric g_{ij} [3] that is

$$g_{ij}(\theta) = \mathbb{E}_{p_\theta}(l_i l_j) = \frac{\partial^2 \psi(\theta)}{\partial \theta_i \partial \theta_j}(\theta),$$

which is positive definite by the maximum rank hypothesis on h . So (\mathcal{E}, g) is a Riemannian manifold. To show that \mathcal{E} is dually flat manifold one needs to introduce Amari α -connection with Christoffel symbols

$$\Gamma_{ij,k}^{(\alpha)} = \mathbb{E}[(\partial_i \partial_j l + \frac{1-\alpha}{2} \partial_i l \partial_j l) \partial_k l], \quad \alpha \in [0, 1]$$

where $l = \ln p(x, \theta)$. By direct computation one shows that $\Gamma^{(1)}$ (resp. $\Gamma^{(-1)}$) vanish in θ (resp. η) coordinates, so θ and η are global affine charts and the curvature $R^{(1)}$ and $R^{(-1)}$ both vanish, making \mathcal{E} dually flat. Note that a connection can be locally flat (zero curvature) without admitting a global affine coordinate system. For an exponential family manifold the Christoffel symbols vanish on the entire domains of the θ and η coordinates, which are convex open sets, so the flat structure is global. For more details see [3].

4. Controlled Evolution in a Dually Flat Manifold

In this Section we show how the dual coordinate system η, θ can be used to describe the controlled evolution of a point on a dually flat manifold. It is well known that the statistical description of a thermodynamic system is given by an exponential family (5) whose feature functions $h(x) \in \mathbb{R}^n$ are the physical observables of interest for the macroscopic description of the system given in terms of their average values $\eta \in \mathbb{R}^n$. The associated natural parameters $\theta = \theta(\eta)$ are interpreted as generalized inverse temperatures. This is a generalization of the fact that if $h(x) \in \mathbb{R}$ represents the energy of the system in the microscopic phase-space configuration $x \in X$, its average η is the internal energy of the system in contact with an heat bath at the inverse temperature $\theta(\eta)$. The heat bath allows one to control the (inverse) temperature θ of the system. This brief review of equilibrium statistical mechanics terminology motivates the name of (thermal) control or input coordinates for $\theta \in \mathbb{R}^n$ and of effect or output coordinate for $\eta \in \mathbb{R}^n$.

The equilibrium statistical mechanics formalism is an example of an exponential family manifold which is an instance of a dually flat manifold; we will be interested in the following two problems formulated for a dually flat n -dimensional Riemannian manifold (M, g) equipped with a dual coordinate systems (η, θ) :

1) forward or direct problem: given a path in control variables $\theta(t) \in \mathbb{R}^n$ to determine the corresponding evolution $\eta(t) \in \mathbb{R}^n$ in effect space;

2) backward or inverse problem: given an observed evolution $\eta(t)$ in effect space, to determine the corresponding control or cause path $\theta(t)$.

Since η and θ are local affine coordinates of a dually flat manifold, the above problems have a solution using (3)

$$\text{direct problem :} \quad \text{given } \theta(t), \text{ the evolution is } \eta(t) = \nabla \psi(\theta(t)) \quad (7)$$

and

$$\text{inverse problem :} \quad \text{given } \eta(t), \text{ the evolution is } \theta(t) = \nabla \varphi(\eta(t)). \quad (8)$$

In the following we will consider two important classes of direct and inverse problems.

Direct Problem: System in a Heat Bath

In [1] the following instance of a direct problem for a control path referred to a k -cut of a mutually dual foliation (see Section 2.1) is considered: let $t \mapsto (\theta_A(t), \bar{\theta}_B)$ be given, that is the system is contained

in a heath bath which realize the constraint $\theta_B(t) = \bar{\theta}_B$. Using (7) we have, adopting for simplicity the shorthand block notation $\psi_{AB} = \partial_{AB}\psi$

$$\begin{cases} \dot{\eta}^A(t) = \frac{d}{dt}(\partial_A\psi(\theta_A(t), \bar{\theta}_B)) = \partial_{AA}\psi \dot{\theta}_A = \psi_{AA}\dot{\theta}_A, \\ \dot{\eta}^B(t) = \frac{d}{dt}(\partial_B\psi(\theta_A(t), \bar{\theta}_B)) = \psi_{BA}\dot{\theta}_A \end{cases}$$

and since $\psi_{AA} \in Mat(k)$ is invertible, we can write $\dot{\theta}_A = \psi_{AA}^{-1}\dot{\eta}^A$ and obtain the result

$$\dot{\eta}^B = \psi_{BA}\psi_{AA}^{-1}\dot{\eta}^A = \Gamma_{BA}(\theta_A(t), \bar{\theta}_B)\dot{\eta}^A, \quad \dot{\eta}^A = \psi_{AA}\dot{\theta}_A. \quad (9)$$

So, if the system is contained in a B -type heath bath $\theta_B(t) = \bar{\theta}_B$, then the evolution in the η^B effect variables can be reconstructed from the evolution of the η^A variables using (9)₁.

This result has an interesting geometrical interpretation: if we consider the local fibration $\eta \circ \pi_A(p) = \eta^A$, with fibres $E(\bar{\eta}^A)$ see (4), the orthogonal submanifold $Q(\bar{\theta}_B)$ defines an orthogonal complement $H(\eta \circ \pi_A)$ to the tangent space to the fibres $V(\eta \circ \pi_A) = \ker T(\eta \circ \pi_A)$ and hence $Q(\bar{\theta}_B)$ is an integrable orthogonal Ehresmann connection, i.e a section of the fibre map $\eta \circ \pi_A$ and the Γ_{BA} are the coefficients of the Ehresmann connection, see e.g. [11], [12]. Therefore (9) can be seen as the horizontal lift of the path $\eta_A(t)$ in base space

$$\dot{\eta}^B = \text{hor}(\eta^A(t)) = \Gamma_{BA}(\theta(t))\dot{\eta}^A(t) \otimes \partial^B$$

where $\text{hor} : \mathbb{R}^k \rightarrow H(\eta \circ \pi_A) = TQ(\theta_B) \subset T_pM$ is the horizontal lift operator.

In the case of an exponential family manifold, the connection coefficients Γ_{BA} have an important interpretation in statistical theory

$$\Gamma_{BA}(\theta) = \psi_{BA}\psi_{AA}^{-1} = \text{cov}(h_A, h_B)\text{cov}(h_A, h_A)^{-1}$$

where $\text{cov}(h_A, h_B)$ is the matrices of the covariances between h_A and h_B observables and $\text{cov}(h_A, h_A)$ is the square matrix of the variance and covariance between h_A type observables. This shows that when the θ_B parameter are kept constant, the inference on the η^B variables given the evolution of the η^A variables is the linear predictor of the B random variables in terms of the A random variables.

The dual of the above system evolution is given by the following:

Inverse Problem: System Insensitive to the Control in the A Variables

Given an evolution in the effect space of the type $\eta(t) = (\bar{\eta}^A, \eta^B(t))$, to determine the corresponding time evolution in the cause parameter $\theta(t)$. Using the same argument as above we have from (7)

$$\begin{cases} \dot{\eta}^A(t) = \frac{d}{dt}(\partial_A\psi(\theta_A(t), \theta_B(t))) = \psi_{AA}\dot{\theta}_A + \psi_{AB}\dot{\theta}_B = 0, \\ \dot{\eta}^B(t) = \frac{d}{dt}(\partial_B\psi(\theta_A(t), \theta_B(t))) = \psi_{BA}\dot{\theta}_A + \psi_{BB}\dot{\theta}_B. \end{cases}$$

Therefore the constraint in effect space induces the following constraint in control space

$$\dot{\theta}_A = -\psi_{AA}^{-1}\psi_{AB}\dot{\theta}_B = -\Gamma_{AB}^T(\theta_A(t), \theta_B(t))\dot{\theta}_B \quad (10)$$

where Γ_{AB}^T is the transpose of the matrix Γ_{BA} . Note that if $\dot{\theta}_B(t) \neq 0$ the above equation (10) can be given the simpler form of an ODE for the function $\theta_A = \theta_A(\theta_B)$ as

$$\frac{d\theta_A}{d\theta_B} = -\Gamma_{AB}^T(\theta_A, \theta_B).$$

For the sake of completeness we derive the above result (10) using the inverse problem formulation in (8). Adopting the shorthand notation $\partial^{AB}\varphi = \varphi^{AB}$ we have

$$\begin{cases} \dot{\theta}_A(t) = \frac{d}{dt}(\partial^A \varphi(\bar{\eta}^A(t), \eta^B)) = \partial^{AB} \varphi \dot{\eta}^B = \varphi^{AB} \dot{\eta}^B, \\ \dot{\theta}_B(t) = \frac{d}{dt}(\partial^B \varphi(\bar{\eta}^A(t), \eta^B)) = \varphi^{BB} \dot{\eta}^B \end{cases}$$

and since $\varphi^{BB} \in \text{Mat}(n-k)$ is invertible, we have that

$$\dot{\theta}_A = \varphi^{AB} (\varphi^{BB})^{-1} \dot{\theta}_B = \Gamma^{AB} \dot{\theta}_B$$

which coincides with the above result (10) because

$$\Gamma^{AB} = \varphi^{AB} (\varphi^{BB})^{-1} = -\psi_{AA}^{-1} \psi_{AB} = -\Gamma_{AB}^T.$$

This last equality can be easily checked by expanding in block representation the identity

$$g^{ij} g_{jk} = \varphi^{ij} \psi_{jk} = \delta_k^i.$$

Example: Direct Problem for the Gaussian Exponential Manifold

Before turning to the mixed control problem we give a simple example of the theory developed so far. It is well known that the gaussian probability density is the maximal entropy density under the constraint of fixed value for the first two moments. Therefore, we can consider for $x \in X = \mathbb{R}$ the feature functions $h_1(x) = x$ and $h_2(x) = x^2$. The probability density defines an exponential family of the form

$$p(\theta_1, \theta_2, x) = e^{\theta_1 x + \theta_2 x^2 - \psi(\theta_1, \theta_2)}$$

where

$$\psi(\theta_1, \theta_2) = \ln \int_{\mathbb{R}} e^{\theta_1 x + \theta_2 x^2} dx = -\frac{\theta_1^2}{4\theta_2} + \frac{1}{2} \ln(-\theta_2) + \frac{1}{2} \ln \pi, \quad \theta_2 < 0.$$

We consider the following forward control problem: to determine the evolution of $\eta^B = \eta^2 = \mathbb{E}_{p_\theta}[x^2]$ given the evolution $\eta^A = \eta^1 = \mathbb{E}_{p_\theta}[x]$ under the control path $(\theta_A(t), \bar{\theta}_B) = (\theta_1(t), \bar{\theta}_2)$. This can be done using (9) by computing the Ehresmann connection coefficient

$$\Gamma_{BA} = \Gamma_{21} = \frac{\psi_{12}}{\psi_{11}} = \frac{\psi_{\theta_1 \theta_2}}{\psi_{\theta_1 \theta_1}} = -\frac{\theta_1}{\theta_2} = -\frac{\theta_1(t)}{\bar{\theta}_2}$$

therefore

$$\dot{\eta}^2(t) = \Gamma_{21}(t) \dot{\eta}^1(t) = -\frac{\theta_1(t)}{\bar{\theta}_2} \dot{\eta}^1(t) \quad (11)$$

which can directly integrated

$$\eta^2(t) = -\frac{1}{\bar{\theta}_2} \int \theta_1(t) \dot{\eta}^1(t) dt + \text{const.}$$

If $\dot{\eta}^1(t) \neq 0$, $\eta^1(t)$ can be locally inverted in $t = t(\eta^1)$ and we can express the horizontal lift (11) in the form

$$\frac{d\eta^2}{d\eta^1} = \Gamma_{21}(t) = \Gamma_{21}(t(\eta^1))$$

so that

$$\eta^2 = \int \Gamma_{21}(t(\eta^1)) d\eta^1 = -\frac{1}{\bar{\theta}_2} \int \theta_1(t(\eta^1)) d\eta^1 + \text{const.}$$

As an example, consider the evolution $\eta^1(t) = t$, with inverse $t(\eta^1) = \eta^1$. We want to determine the control path $\theta_1(t(\eta^1)) = \theta_1(\eta^1)$ so that the variance is a given function $\text{var}(h_1) = \eta^2 - (\eta^1)^2 = f(\eta^1)$. We have the functional equation

$$f(\eta^1) = \eta^2 - (\eta^1)^2 = -\frac{1}{\theta_2} \int \theta_1(\eta^1) d\eta^1 + \text{const} - (\eta^1)^2$$

which gives by derivation

$$\theta_1(\eta^1) = (-\bar{\theta}_2)(2\eta^1 + f'(\eta^1)).$$

5. Mixed Control Problem and Inferential Scattering

The notion of Inferential scattering was exposed in a couple of papers by E.T. Jaynes [1,2]. Jaynes' original exposition refers to equilibrium statistical mechanics. In this setting, the results of the above Section describe the inference on the evolution of the average value of a group of random variables (B) given the evolution of the average value of another group (A) of random variables under the hypothesis that the random variables are described by an exponential family and a constraint is in force. It can be seen as an inference result.

To generalize the notion of inferential scattering we need to introduce the notion of (k, l) -cut in a n -dimensional dual coordinate system of a dually flat manifold: let (l, k) with $1 < l < k < n$ and consider the splitting of the coordinates

$$\eta = (\eta^a, \eta^c, \eta^B), \quad \theta = (\theta_a, \theta_c, \theta_B), \quad a = 1, \dots, l, \quad c = l + 1, \dots, k, \quad B = k + 1, \dots, n.$$

In practice, we split the A block variables in $A = (a, c)$. Now we can consider the following instance of mixed control problem

Mixed Control Problem for a System in a B-Type Heat Bath and Insensitive to the Control in the c Variables

This is a combination of the two above cases. We want to determine the evolution in the η^B variables under the controlled evolution $\theta_a(t)$ knowing that we have the constraint $\eta^c(t) = \bar{\eta}^c$ in the effect space and $\theta_B(t) = \bar{\theta}_B$ in cause space. Reasoning as above we consider the system of evolution equations

$$\begin{cases} \dot{\eta}^a = \frac{d}{dt}(\partial_a \psi(\theta_a(t), \theta_c(t), \bar{\theta}_B)) = \psi_{aa} \dot{\theta}_a + \psi_{ac} \dot{\theta}_c, \\ \dot{\eta}^c = \frac{d}{dt}(\partial_c \psi(\theta_a(t), \theta_c(t), \bar{\theta}_B)) = \psi_{ca} \dot{\theta}_a + \psi_{cc} \dot{\theta}_c = 0, \\ \dot{\eta}^B = \frac{d}{dt}(\partial_B \psi(\theta_a(t), \theta_c(t), \bar{\theta}_B)) = \psi_{Ba} \dot{\theta}_a + \psi_{Bc} \dot{\theta}_c. \end{cases}$$

From the second equation we get

$$\dot{\theta}_c = -\psi_{cc}^{-1} \psi_{ca} \dot{\theta}_a \quad (12)$$

and we substitute $\dot{\theta}_c$ in the first and third equation to obtain

$$\begin{cases} \dot{\eta}^a = \psi_{aa} \dot{\theta}_a - \psi_{ac} \psi_{cc}^{-1} \psi_{ca} \dot{\theta}_a = \psi_{aa} (\mathbb{I} - \psi_{aa}^{-1} \psi_{ac} \psi_{cc}^{-1} \psi_{ca}) \dot{\theta}_a = \psi_{aa} (\mathbb{I} - R(a, c)) \dot{\theta}_a \\ \dot{\eta}^B = \psi_{Ba} \dot{\theta}_a - \psi_{Bc} \psi_{cc}^{-1} \psi_{ca} \dot{\theta}_a = (\psi_{Ba} - \psi_{Bc} \psi_{cc}^{-1} \psi_{ca}) \dot{\theta}_a \end{cases} \quad (13)$$

where \mathbb{I} is the identity matrix and we have introduced the correlation coefficient matrix

$$R(a, c) = \psi_{aa}^{-1} \psi_{ac} \psi_{cc}^{-1} \psi_{ca}. \quad (14)$$

The last equality in (13)₁ can be inverted

$$\dot{\theta}_a = (\mathbb{I} - R(a, c))^{-1} \psi_{aa}^{-1} \dot{\eta}^a$$

and substituted in (13)₂ to finally obtain

$$\dot{\eta}^B = (\psi_{Ba} - \psi_{Bc}\psi_{cc}^{-1}\psi_{ca})(\mathbb{I} - R(a, c))^{-1}\psi_{aa}^{-1}\dot{\eta}^a = \mathcal{G}_{Ba}\dot{\eta}^a. \quad (15)$$

This is our inference result on the evolution of the η^B variables given the mixed (forward and backward) control. Note that $\mathcal{G}_{Ba} = \mathcal{G}_{Ba}(\theta_a(t), \theta_c(t), \bar{\theta}_B)$ but the evolution in the controlled variables θ_a, θ_c is subject to the equation (12) which enforces the constraint in effect space $\eta^c = \bar{\eta}^c$.

5.1. The Inferential Scattering Formula

Since $R(a, c)$ in (14) is the analogous of the correlation coefficient of two random variables

$$\rho(A, B) = \frac{\text{cov}(A, B)^2}{\text{var}(A)\text{var}(B)}, \quad 0 < \rho(A, B) < 1$$

it reasonable to assume that the matrix $R(a, c)$ has spectral radius less than one. Under this assumption, the following Taylor expansion hold

$$(\mathbb{I} - R(a, c))^{-1} = \mathbb{I} + R(a, c) + R(a, c)^2 + R(a, c)^3 + \dots \quad (16)$$

If we substitute (16) in (15) we readily obtain the following expansion

$$\dot{\eta}^B = (\Gamma_{Ba} - \Gamma_{Bc}\Gamma_{ca} + \Gamma_{Ba}\Gamma_{ac}\Gamma_{ca} - \Gamma_{Bc}\Gamma_{ca}\Gamma_{ac}\Gamma_{ca} + \dots)\dot{\eta}^a \quad (17)$$

where we have set

$$\Gamma_{Ba} = \psi_{Ba}\psi_{aa}^{-1}, \quad \Gamma_{Bc} = \psi_{Bc}\psi_{cc}^{-1}, \quad \Gamma_{ca} = \psi_{ca}\psi_{aa}^{-1}, \quad \Gamma_{ac} = \psi_{ac}\psi_{cc}^{-1}.$$

The above expansion formula (17) has the following interpretation, see [1]: the term $\Gamma_{Ba}d\eta^a$ is the inferred (infinitesimal) variation of the B observable induced by a variation $d\eta^a$ of the a observables. The term $\Gamma_{Bc}\Gamma_{ca}d\eta^a$ is the inferred variation in B as induced by a variation $d\eta_c = \Gamma_{ca}d\eta^a$ of the c observables which is in turn induced by the variation $d\eta^a$. The other terms in (17) have a similar interpretation. This result has been called Inferential Scattering formula [2] by analogy with the Born series expansion of the operator which describes the multiple scattering phenomenon [10]: a wave can reach a point B in space directly from source a (Γ_{Ba} term) or after having been reflected by an obstacle in c ($\Gamma_{Bc}\Gamma_{ca}$ term) or after having been reflected back in a from c and then arriving at B ($\Gamma_{Ba}\Gamma_{ac}\Gamma_{ca}$ term). See diagram (18) below. Note that the first term in (17) Γ_{Ba} coincides with the one derived before in (9) in case of forward control. Of course the size of the higher order correction terms after Γ_{Ba} depends on the norm of the correlation matrix $R(a, c)$.



5.2. A Mixed Control Example: Neural Network of Two Neurons

The following toy model (see e.g. [5]) describes the neural activity of a set of two neurons using two discrete random variables which can take the value 1 (active) or 0 (silent). The state space of the network is $X = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$. Given $x = (x_1, x_2) \in X$, let $p(x)$ be the probability that the network is in the state x . As observables we consider $h_i(x) = x_i$, $i = 1, 2$ which describe the state of the two neurons and $h_3(x) = x_1x_2$. The exponential manifold is comprised of discrete probability distributions of the form

$$p(x) = e^{\theta \cdot h - \psi(\theta)} = e^{\theta_1 x_1 + \theta_2 x_2 + \theta_3 x_1 x_2 - \psi(\theta)}$$

where

$$\psi(\theta) = \ln \left(\sum_{x \in X} e^{\theta_1 x_1 + \theta_2 x_2 + \theta_3 x_1 x_2} \right) = \ln (1 + e^\theta + e_2^\theta + e^{\theta_1 + \theta_2 + \theta_3}) = \ln Z(\theta)$$

The associated effect space coordinates

$$\eta^i = \mathbb{E}_{p_\theta} [x_i] = \frac{\partial \psi}{\partial \theta_i}, \quad i = 1, 2 \quad \eta^3 = \mathbb{E}_{p_\theta} [x_1 x_2] = \frac{\partial \psi}{\partial \theta_3}$$

are called firing rate of neurons $i = 1, 2$ and joint firing rate. A direct computation shows that

$$\eta^i = \frac{e^{\theta_i} + e^{\theta_1 + \theta_2 + \theta_3}}{Z(\theta)} = \text{Prob}(x_i = 1), \quad i = 1, 2, \quad \eta^3 = \frac{e^{\theta_1 + \theta_2 + \theta_3}}{Z(\theta)} = \text{Prob}(x_1 x_2 = 1).$$

The Riemannian metric matrix $g \in \text{Mat}(3)$ can also be computed directly as

$$g_{ij}(\theta) = \frac{\partial^2 \psi(\theta)}{\partial \theta_i \partial \theta_j}.$$

We investigate the mixed control problem (see Section 5) for this system. We consider the $(2, 1)$ -cut in the 3-dimensional dual coordinate system defined by the following identification

$$(\eta^1, \eta^2, \eta^3) = (\eta^a, \eta^c, \eta^B), \quad (\theta_1, \theta_2, \theta_3) = (\theta_a, \theta_c, \theta_B).$$

We want to determine the evolution in the $\eta^B = \eta^3$ variable (joint firing rate of neuron 1 and 2) in terms of the evolution of $\eta^a = \eta^1$ (firing rate of neuron 1) under the control $\theta_B(t) = \theta_3(t)$ knowing that we have the constraint $\eta^c(t) = \eta^2(t) = \bar{\eta}^2$ (firing rate of neuron 2 is insensitive to the control) in the effect space. To this we can use the perturbation expansion (17)

$$d\eta^3 = (\Gamma_{31} - \Gamma_{32}\Gamma_{21} + \Gamma_{31}\Gamma_{12}\Gamma_{21} - \Gamma_{32}\Gamma_{21}\Gamma_{12}\Gamma_{21} + \dots) d\eta^1$$

where

$$\Gamma_{31} = \frac{\psi_{31}}{\psi_{11}}, \quad \Gamma_{32} = \frac{\psi_{32}}{\psi_{22}}, \quad \Gamma_{21} = \frac{\psi_{21}}{\psi_{11}}, \quad \Gamma_{12} = \frac{\psi_{12}}{\psi_{22}}$$

and

$$R(1, 2) = \Gamma_{12}\Gamma_{21} = \frac{\psi_{12}\psi_{21}}{\psi_{11}\psi_{22}} = \frac{\psi_{12}^2}{\psi_{11}\psi_{22}} = \frac{\text{cov}(x_1, x_2)^2}{\text{var}(x_1)\text{var}(x_2)} = \rho^2(x_1, x_2)$$

where $\rho(x_1, x_2) \in [0, 1]$ is Pearson's correlation coefficient which is a measure of the strength of the linear correlation (normalized covariance) between the two variables. So the perturbation parameter in the Born series expansion (16) is the correlation coefficient. The explicit expression (setting $\theta_1 = x$, $\theta_2 = y$ and $\theta_3 = z$ for simplicity) is

$$R(1, 2)(x, y, z) = \left(\frac{1}{1 + e^x} - \frac{1}{e^{x+z}} \right) \left(\frac{1}{1 + e^y} - \frac{1}{e^{y+z}} \right).$$

Note that the two control variables (x, y) are not completely free because relation (12) which enforces the constraint $\eta^3 = \eta^c = \bar{\eta}^c$ holds. In our notation after a straightforward computation (12) reads

$$\dot{y} = -\frac{\psi_{12}}{\psi_{22}} \dot{x} = \left(\frac{1}{1 + e^x} - \frac{1}{1 + e^{x+\bar{z}}} \right) \dot{x} = -\Gamma_{12}(x, \bar{z}) \dot{x}$$

where we have set $z = \theta_3 = \bar{\theta}_3 = \bar{z}$. If $\dot{x} \neq 0$ the above equation can be readily integrated

$$\frac{dy}{dx} = \Gamma_{12}(x, \bar{z}) \quad \text{hence} \quad y(x) = \int \Gamma_{12}(x, \bar{z}) dx + \text{const.}$$

to give

$$y(x, \bar{z}) = 2[\operatorname{artanh}(1 + 2e^{x+\bar{z}}) - \operatorname{artanh}(1 + 2e^x)].$$

Therefore, in the mixed control case, the correlation coefficient $R(1,2)$ is a function of $\theta_1 = x$ only, of the form

$$R_{\bar{z}}(1,2)(x) = R(1,2)(x, y(x, \bar{z}), \bar{z}).$$

A plot of $R(1,2)$ for $\bar{z} = \bar{\theta}_3 = 5$ is in Figure 2. The figure helps to visualize the size of the perturbation expansion parameter $R(1,2)$. This simple example is only suggestive of the possible applications of the notion of Inferential Scattering.

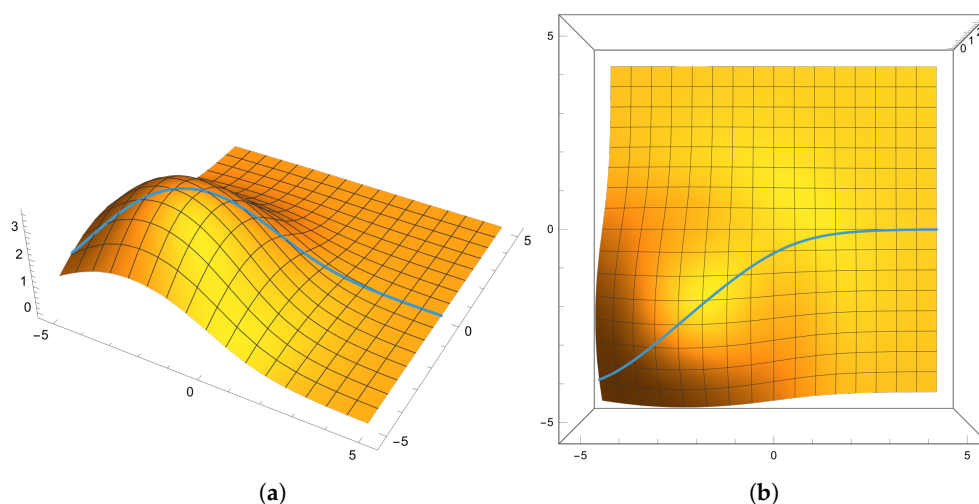


Figure 2. Plot of the correlation coefficient $R(1,2)$ as a function of $x = \theta_1$ and $y = \theta_2$ for $(\theta_1, \theta_2) \in [-5, 5]^2$ and $\bar{z} = \bar{\theta}_3 = 5$ (forward control). The blue line gives the allowed values of $\theta_2 = \theta_2(\theta_1)$ as a consequence of the constraint in effect space (backward control) $\eta^3 = \bar{\eta}^3$. We have plotted $R(1,2) \times 5$ to enhance visibility in the plots. (a) : side view. (b) : top view.

6. Discussion

The primary aim of this paper was to embed the notion of Inferential Scattering as exposed by E.T. Jaynes for a thermodynamical system into the language and the concepts of Information Geometry. The key geometrical object turns out to be a dually flat manifold with its coupled bi-orthogonal coordinate systems η and θ . These coordinates have a simple and appealing interpretation in the statistical mechanics framework as (thermal) control and effect space coordinates.

We have also shown that the forward and backward control problems can be seen as a horizontal lift (parallel transport) with respect to an integrable orthogonal Ehresmann connection. In this way the controlled evolution of the system in effect space is globally defined providing a generalization of the original formulation for infinitesimal displacements contained in the Jaynes papers. We stress that this notion of controlled evolution has a potentially wide application range being a form of inference, hence applicable to every statistical system with insufficient information. The geometrical analysis presented in this paper could be further investigated and we plan to do it in a subsequent paper: for example, it remains to be seen its formulation in the larger setting of Hessian manifolds and the possible developments using a weaker setting for the Legendre transform. We hope that the material presented here can renew interest in the scientific community for the original contribution of E.T. Jaynes and its potential applications.

Funding: This research received no external funding.

Acknowledgments: In this section you can acknowledge any support given which is not covered by the author contribution or funding sections.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Jaynes, E.T. Generalized scattering. In Ray Smith, C. & Grandy, W.T. Jr. (Eds), *Maximum-Entropy and Bayesian Methods in Inverse Problems* Dordrecht: D. Reidel Publishing Company, 1985, 377–398.
2. Jaynes, E.T. *Inferential scattering*, 1993. Available online: <https://bayes.wustl.edu/etj/articles/cinfscat.pdf> (accessed on 03/03/2026)
3. Amari, S. & Nagaoka, H. . *Methods of Information Geometry*. (D. Harada, trans.) AMS Translations of Mathematical Monographs, 191, New York: AMS and Oxford University Press, 2000
4. Amari, S. I. . Information geometry of multiple spike trains. In *Analysis of Parallel Spike trains* . Boston, MA: Springer US, 2010, 221-252
5. Amari, S. *Information geometry and its applications*, Springer, 2016.
6. Amari, S. I. . Information geometry on hierarchy of probability distributions. *IEEE transactions on information theory*, **2002**, 47(5), 1701-1711.
7. Amari, S. Information geometry and manifolds of neural networks, *From Statistical Physics to Statistical Inference and Back*. Dordrecht: Springer Netherlands, 1994, 113-138.
8. Amari, S., Kurata, K., Nagaoka, H. Information Geometry of Boltzmann machines *IEEE Trans. on Neural Networks*, **1992**, 3(2), 260–271.
9. Nielsen, F. An elementary introduction to information geometry, *Entropy*, **2020** 22.10 : 1100.
10. Newton, Roger G. *Scattering theory of waves and particles* Springer Science & Business Media, 2013.
11. Marsden, Jerrold E., Montgomery R., and Tudor S. Rațiu. *Reduction, symmetry, and phases in mechanics*. Vol. 436. American Mathematical Soc., 1990.
12. Favretti, M. Geometry and control of thermodynamic systems described by generalized exponential families, *Journal of Geometry and Physics*, **2022**, 176, 104497.

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.