



Conference Proceedings Paper – Entropy

Information geometry in Gaussian random fields: searching for an arrow of time in complex systems

Alexandre Levada ^{1,*}

¹ Departamento de Computação, Universidade Federal de São Carlos, Rod. Washington Luis, Km. 235, São Carlos, Brasil

Received: 24 August 2014 / Accepted: 22 October 2014 / Published: 3 November 2014

Abstract: Random fields are characterized by intricate non-linear relationships between their elements over time. However, what is a reasonable intrinsic definition for time in such complex systems? Here, we discuss the problem of characterizing the notion of time in isotropic pairwise Gaussian random fields. In particular, we are interested in studying the behavior of these fields when temperature deviates from infinity. Our investigations are focused in the relation between entropy and Fisher information, by the definition of the Fisher curve. The results suggest the emergence of an arrow of time as a consequence of asymmetrical geometric deformations in the random field model's metric tensor. In terms of information geometry, the process of taking a random field from a lower entropy state A to a higher entropy state B and then bringing it back to A, induces a natural intrinsic one-way direction of evolution. In practice, there are different trajectories in the information space, suggesting that the deformations induced by the metric tensor into the parametric space (manifold) are not reversible for positive and negative displacements in the inverse temperature parameter direction. In other words, there is only one possible orientation to move through different entropic states along a Fisher curve.

Keywords: information geometry; random fields; arrow of time; Fisher information; entropy

1. Introduction

Since the origins of the human race, the concept of time has always intrigued mankind. Along centuries of evolution many philosophers and researchers have studied this fascinating but seemingly obscure topic [1–3]. What is time? Does it really exist? Why does time seem to flow in one single direction? Is the passage of time merely an illusion? We certainly do not have definitive answers to all these questions. In an attempt to study the effect of the passage of time in complex systems, this paper proposes to investigate a reasonable way to characterize an intrinsic notion of time in random fields composed by Gaussian variables. Our study focuses on an information-theoretic perspective, motivated by the connection between Fisher information and the geometric structure of stochastic models, provided by information geometry [4,5]. The proposed framework is mostly based on a data-driven approach, that is, we make use of intensive computational simulations to achieve our conclusions.

During the last decades, the notion of information has increasingly become more present and relevant in any scale of modern society, as the volume of data that is being produced by scientific experiments is larger than ever. Being able to decode the symbols in this ocean of data is an essential step to learn, understand and assess the rules governing complex phenomena that are part of our world. A major challenge in dealing with complex systems concerns the mining, identification and further classification of patterns and symbols that convey relevant information about the underlying processes that govern their behavior. After the pieces of information are gathered, the global shape starts to emerge, as if an intricate puzzle had been solved. In this scenario, computational tools for exploratory data analysis are a fundamental component of this data-driven knowledge discovery process.

However, one drawback of intensive data-driven approaches is that most exploratory data analysis methods usually rely on the independence assumption, that is, there is no relation between any two random variables, making it difficult to quantify the influence of a set of variables over another set of variables in a random sample from a parametric stochastic model. Independent and identically distributed random variables belonging to the exponential family of probability distributions define the basic approach for any classical statistical inference framework [6–8].

Random field models arise as a natural generalization of the classical approach by the simple replacement of the independence assumption by a more realistic conditional independence hypothesis [9,10]. Basically, in a Markov random field (MRF), knowledge of a finite-support neighborhood around a given variable isolates it from all the remaining variables. A further simplification consists in considering a pairwise interaction model, which means that we are constraining the size of the maximum clique to be two. In other words, a pairwise model captures only binary relationships. Furthermore, if the random field model is isotropic, all the information regarding the spatial dependence structure of the system is conveyed by a single coupling parameter, from now on denoted by β . This parameter is widely known as the inverse temperature of the system. As the value of this parameter deviates from zero, the more our model deviates from the classical statistical scenario (regular exponential family + independence hypothesis). In the Gaussian case for instance, by introducing some degree of dependence between the random variables, assuming that $\beta \neq 0$, we are essentially moving towards a curved exponential family.

Basically, the main goal of this investigation is to use information geometry as a mathematical tool to measure the emergence of an intrinsic notion of time in complex systems modeled by random fields in

which temperature is allowed to deviate from infinity. Computational simulations validate our claim that the arrow of time is possibly a consequence of asymmetrical geometric deformations in the metric tensor (Fisher information) of the statistical manifold of the random field model, when the inverse temperature parameter is disturbed.

Recently, expressions to compute the expected Fisher information regarding the inverse temperature parameter in Gaussian Markov random field models have been proposed by the authors [11]. Also, Markov Chain Monte Carlo simulations have shown that expressing Fisher information in terms of tensor (Kronecker) and pointwise (Hadamard) matrix products leads to more efficient and faster computations. Finally, an indirect but fundamental problem involved in the measurement of these information-theoretic quantities is the estimation of the inverse temperature parameter of a random field, given the observation of a single snapshot of the system. More details about the proposed methodology and the obtained results are discussed in later sections of the paper.

2. Fisher Information in MRF models

The concept of Fisher information [12,13] has been shown to reveal several properties of statistical procedures, from lower bounds on estimation methods [6–8] to information geometry [4,5]. In summary, we can think of Fisher information as being a likelihood analog of Shannon entropy, which is a probability-based measure of uncertainty.

In this paper, our goal is to show that, in an exploratory data analysis context, Fisher information plays a central role in providing tools for measuring and quantifying the behavior of random fields. In this scenario, the most interesting feature of random field models over the classic statistical ones is the possibility to take the dependence between pieces of information into account. Moreover, this underlying dependence structure arises in terms of the system’s temperature, which may even be variable along time.

2.1. The Information Equality

It is known from statistical inference theory that information equality holds for independent observations in the regular exponential family of distributions [6–8]. In other words, we can compute the expected Fisher information of a statistical model $p(X|\theta)$ regarding a parameter of interest θ by two equivalent ways, since it is possible to interchange the integration and differentiation operators:

$$I(\theta; X) = E \left[\left(\frac{\partial}{\partial \theta} \log l(\theta; X) \right)^2 \right] = -E \left[\frac{\partial^2}{\partial \theta^2} \log l(\theta; X) \right] \quad (1)$$

where $l(\theta; X)$ denotes the likelihood function, that is, the probability density function (pdf) interpreted as a function of the model parameters. In our investigations, we replace the pdf of the model by a local conditional density function, which according to the Hammersley-Clifford theorem [14], characterizes the random field model. In fact, this remarkable result states the equivalence between Markov random Fields (local model) and Gibbs random fields (global model).

However, given the intrinsic spatial dependence structure of random field models, information equality is not a natural condition. In general, when the inverse temperature parameter gradually drifts apart from zero (T deviates from infinity), this information “equilibrium” fails. Thus, in random field

models we have to consider two different versions of Fisher information, from now on denoted by type-I (due to the first derivative of the likelihood function) and type-II (due to the second derivative of the likelihood function). Eventually, when certain conditions are satisfied, these two values of information will converge to a unique bound. Therefore, in random fields, these two versions of Fisher information play distinct roles, especially in quantifying the uncertainty in the estimation of the inverse temperature parameter.

2.2. Fisher Information in the Gaussian Markov Random Field Model

Gaussian random fields are important models for dealing with spatially dependent continuous random variables once they provide a general framework for studying the non-linear interactions between elements of a stochastic complex system along time. One of the major advantages of these models is the mathematical tractability, which allows us to derive exact closed-form expressions for both maximum pseudo-likelihood estimators of the inverse temperature parameter and expected Fisher information. Due to the Hammersley-Clifford theorem, it is also possible to characterize these random fields by a set of local conditional density functions (LCDF's), avoiding the use of the joint Gibbs distribution.

Definition 1. *An isotropic pairwise Gaussian Markov random field regarding a local neighborhood system η_i defined on a lattice $S = \{s_1, s_2, \dots, s_n\}$ is completely characterized by a set of n local conditional density functions $p(x_i|\eta_i, \vec{\theta})$, given by:*

$$p(x_i|\eta_i, \vec{\theta}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{1}{2\sigma^2} \left[x_i - \mu - \beta \sum_{j \in \eta_i} (x_j - \mu) \right]^2 \right\} \quad (2)$$

with $\vec{\theta} = (\mu, \sigma^2, \beta)$, where μ and σ^2 are the expected value and the variance of the random variables, and β is the inverse temperature or coupling parameter. Note that for $\beta = 0$, the model degenerates to the usual Gaussian distribution, which belongs to the regular exponential family. It has been shown that the geometric structure of regular exponential family distributions exhibit constant curvature. It is also known that from an information geometry perspective [4,5], the natural Riemannian metric of these probability distribution manifolds is given by the Fisher information. However, little is known about information geometry on more general statistical models, such as random field models. In this paper, our primary objective is to study, from an information theory perspective, how changes in the inverse temperature parameter affect the metric tensor of Gaussian Markov random field models, more precisely, the Fisher information regarding the β parameter. By measuring this quantity we are actually capturing and quantifying an important component of a complex deformation process induced by a displacement in the inverse temperature parameter direction.

2.3. Maximum Pseudo-Likelihood Estimation in MRF Models

Before we can compute the expected Fisher information in a random field, it is necessary to estimate the model parameters. In this paper, the Gaussian Markov random field parameters μ and σ^2 are both estimated by the sample mean and variance, respectively (the maximum likelihood estimatives).

However, maximum likelihood estimation is intractable for the inverse temperature parameter estimation, due to the existence of the partition function in the joint Gibbs distribution. An alternative, proposed by Besag [15], is to perform maximum pseudo-likelihood estimation, which is based on the conditional independence principle. The pseudo-likelihood function is defined as the product of the LCDF's for all the observations in the random field.

Definition 2. Let an isotropic pairwise Markov random field model be defined on a rectangular lattice $S = \{s_1, s_2, \dots, s_n\}$ with a neighborhood system η_i . Assuming that $\mathbf{X}^{(t)} = \{x_1^{(t)}, x_2^{(t)}, \dots, x_n^{(t)}\}$ denotes the set corresponding to the observations at a time t (a snapshot of the random field), the pseudo-likelihood function of the model is defined by:

$$L(\beta; \mathbf{X}^{(t)}) = \prod_{i=1}^n p(x_i | \eta_i, \beta) \quad (3)$$

The pseudo-likelihood function is the product of the local conditional density functions throughout the field. Note that the pseudo-likelihood function is a function of the model parameters.

2.4. Estimating the Inverse Temperature in the GMRF Model

To derive the maximum pseudo-likelihood (MPL) estimator of the inverse temperature parameter we proceed by plugging equation (2) into equation (3) and taking the logarithm:

$$\log L(\vec{\theta}; \mathbf{X}^{(t)}) = -\frac{n}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n \left[x_i - \mu - \beta \sum_{j \in \eta_i} (x_j - \mu) \right]^2 \quad (4)$$

By differentiating equation (4) with respect to β and properly solving the pseudo-likelihood equation we obtain the following MPL estimator for β :

$$\hat{\beta}_{MPL} = \frac{\sum_{i=1}^n \left[(x_i - \mu) \sum_{j \in \eta_i} (x_j - \mu) \right]}{\sum_{i=1}^n \left[\sum_{j \in \eta_i} (x_j - \mu) \right]^2} \quad (5)$$

Considering that the model is defined on a regular rectangular 2D lattice, the cardinality of the neighborhood system, $|\eta_i|$, is spatially invariant. Thus each variable is dependent on a fixed number of neighbors in the lattice and the maximum pseudo-likelihood estimator for the inverse temperature can be rewritten in terms of cross covariances, since equation (5) can be expressed as:

$$\hat{\beta}_{MPL} = \frac{\sum_{j \in \eta_i} \hat{\sigma}_{ij}}{\sum_{j \in \eta_i} \sum_{k \in \eta_i} \hat{\sigma}_{jk}} \quad (6)$$

where $\hat{\sigma}_{ij}$ denotes the sample covariance between the central variable x_i and $x_j \in \eta_i$. Similarly, $\hat{\sigma}_{jk}$ denotes the sample covariance between two variables belonging to the neighborhood system η_i (note

that the definition of a neighborhood system η_i does not include the the location s_i). All the information regarding the inverse temperature parameter is conveyed by the covariance matrix of the observable local interaction patterns (second-order statistics), which is somehow expected since we are dealing with Gaussian random variables.

2.5. Expected Fisher Information

As mentioned earlier, in the GMRF model it is possible to obtain exact closed-form expressions for the expected Fisher information. In this section, we proceed with the derivation of both type-I (Φ_β) and type-II (Ψ_β) expected Fisher information in isotropic pairwise Gaussian Markov random fields. In recent research efforts, the authors have already derived these expressions in a previous investigation. In this paper, we will not cover all the steps involved in the derivation. For more details, the reader is referred to [11]. To derive an expression for Φ_β , we plug the LCDf of the isotropic pairwise GMRF model (equation 2) in the first definition of Fisher information in equation (1) to obtain the following:

$$\begin{aligned}
\Phi_\beta &= \frac{1}{\sigma^4} E \left\{ \left[\left(x_i - \mu - \beta \sum_{j \in \eta_i} (x_j - \mu) \right) \left(\sum_{j \in \eta_i} (x_j - \mu) \right) \right]^2 \right\} \\
&= \frac{1}{\sigma^4} E \left\{ \left[(x_i - \mu)^2 - 2\beta \sum_{j \in \eta_i} (x_i - \mu) (x_j - \mu) + \beta^2 \sum_{j \in \eta_i} \sum_{k \in \eta_i} (x_j - \mu) (x_k - \mu) \right] \times \right. \\
&\quad \left. \left[\sum_{j \in \eta_i} \sum_{k \in \eta_i} (x_j - \mu) (x_k - \mu) \right] \right\} \\
&= \frac{1}{\sigma^4} \left\{ \sum_{j \in \eta_i} \sum_{k \in \eta_i} E [(x_i - \mu) (x_i - \mu) (x_j - \mu) (x_k - \mu)] \right. \\
&\quad - 2\beta \sum_{j \in \eta_i} \sum_{k \in \eta_i} \sum_{l \in \eta_i} E [(x_i - \mu) (x_j - \mu) (x_k - \mu) (x_l - \mu)] \\
&\quad \left. + \beta^2 \sum_{j \in \eta_i} \sum_{k \in \eta_i} \sum_{l \in \eta_i} \sum_{m \in \eta_i} E [(x_j - \mu) (x_k - \mu) (x_l - \mu) (x_m - \mu)] \right\}
\end{aligned} \tag{7}$$

Using the Isserlis' theorem [16] to compute higher-order moments of normally distributed random variables we have:

$$\begin{aligned}
\Phi_\beta &= \frac{1}{\sigma^4} \left\{ \sum_{j \in \eta_i} \sum_{k \in \eta_i} [\sigma^2 \sigma_{jk} + 2\sigma_{ij} \sigma_{ik}] - 2\beta \sum_{j \in \eta_i} \sum_{k \in \eta_i} \sum_{l \in \eta_i} [\sigma_{ij} \sigma_{kl} + \sigma_{ik} \sigma_{jl} + \sigma_{il} \sigma_{jk}] \right. \\
&\quad \left. + \beta^2 \sum_{j \in \eta_i} \sum_{k \in \eta_i} \sum_{l \in \eta_i} \sum_{m \in \eta_i} [\sigma_{jk} \sigma_{lm} + \sigma_{jl} \sigma_{km} + \sigma_{jm} \sigma_{kl}] \right\}
\end{aligned} \tag{8}$$

where σ_{ij} is the covariance between the central variable x_i and a neighbor $x_j \in \eta_i$ and σ_{jk} , σ_{kl} , σ_{jl} , σ_{lm} , σ_{km} and σ_{jm} are covariances between two neighboring variables in η_i . Therefore, we can express Φ_β

in terms of the covariances between the random variables in a local neighborhood system, which means that we can use the covariance matrix of the local patterns to compute it. Following the same approach, it is possible to derive an expression to the type-II expected Fisher information, Ψ_β , as:

$$\begin{aligned}\Psi_\beta &= \frac{1}{\sigma^2} E \left\{ \left[\sum_{x_j \in \eta_i} (x_j - \mu) \right]^2 \right\} = \frac{1}{\sigma^2} E \left[\sum_{x_j \in \eta_i} \sum_{x_k \in \eta_i} (x_j - \mu) (x_k - \mu) \right] \\ &= \frac{1}{\sigma^2} \left\{ \sum_{x_j \in \eta_i} \sum_{x_k \in \eta_i} E [(x_j - \mu) (x_k - \mu)] \right\} = \frac{1}{\sigma^2} \sum_{j \in \eta_i} \sum_{k \in \eta_i} \sigma_{jk}\end{aligned}\quad (9)$$

where σ_{jk} is the covariance between two neighboring variables in η_i . Note that unlike Φ_β , Ψ_β does not depend explicitly on β (inverse temperature).

In order to simplify the notations and also to make computations faster, the expressions for Φ_β and Ψ_β can be rewritten in a matrix-vector form. Let Σ_p be the covariance matrix of the random vectors $\vec{p}_i, i = 1, 2, \dots, n$, obtained by lexicographic ordering the local configuration patterns $x_i \cup \eta_i$. In this work, we choose a second-order neighborhood system, making each local configuration pattern a 3×3 patch. Thus, since each vector \vec{p}_i has 9 dimensions, the resulting covariance matrix Σ_p is 9×9 . Let Σ_p^- be the sub-matrix of dimensions 8×8 obtained by removing the central row and central column of Σ_p (these elements are the covariances between x_i and each one of its neighbors $x_j \in \eta_i$). Also, let $\vec{\rho}$ be the vector of dimensions 8×1 formed by all the elements of the central row of Σ_p , excluding the middle one (which is the variance of x_i actually). Figure 1 illustrates the process of decomposing the covariance matrix Σ_p into the sub-matrix Σ_p^- and the vector $\vec{\rho}$ in an isotropic pairwise GMRF model defined on a second-order neighborhood system (8 nearest neighbors).

Given the above, we can express equations (8) and (9) in a tensorial form using Kronecker products. The following definitions provide a computationally efficient way to compute both Φ_β and Ψ_β exploring tensor products.

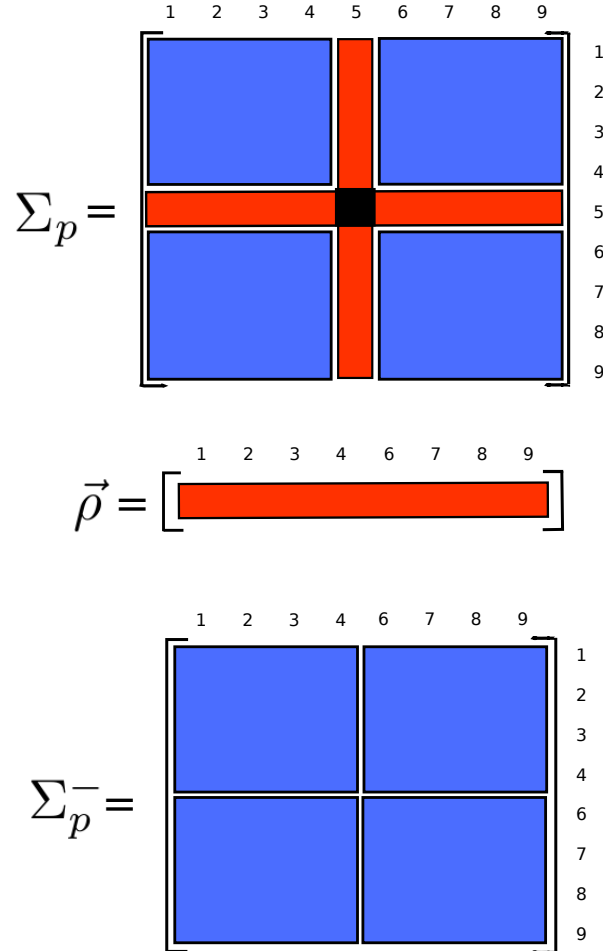
Definition 3. Let an isotropic pairwise GMRF be defined on a lattice $S = \{s_1, s_2, \dots, s_n\}$ with a neighborhood system η_i of size K (usual choices for K are even values: 4, 8, 12, 20 or 24). Assuming that $\mathbf{X}^{(t)} = \{x_1^{(t)}, x_2^{(t)}, \dots, x_n^{(t)}\}$ denotes the global configuration of the system at time t , and both $\vec{\rho}$ and Σ_p^- are defined according to Figure 1, the type-I expected Fisher information Φ_β for $\mathbf{X}^{(t)}$ is:

$$\Phi_\beta = \frac{1}{\sigma^2} \|\Sigma_p^-\|_+ + \frac{1}{\sigma^4} \left[2 \|\vec{\rho} \otimes \vec{\rho}^T\|_+ - 6\beta \|\vec{\rho}^T \otimes \Sigma_p^-\|_+ + 3\beta^2 \|\Sigma_p^- \otimes \Sigma_p^-\|_+ \right] \quad (10)$$

where $\|A\|_+$ denotes the summation of all the entries of the matrix A (not to be confused with a matrix norm) and \otimes denotes the Kronecker (tensor) product. Similarly, it is possible to define Ψ_β using a matrix-vector notation and tensor products.

Definition 4. Let an isotropic pairwise GMRF be defined on a lattice $S = \{s_1, s_2, \dots, s_n\}$ with a neighborhood system η_i of size K (usual choices for K are 4, 8, 12, 20 or 24). Assuming that $\mathbf{X}^{(t)} = \{x_1^{(t)}, x_2^{(t)}, \dots, x_n^{(t)}\}$ denotes the global configuration of the system at time t and Σ_p^- is defined according Figure 1, the type-II expected Fisher information Ψ_β for $\mathbf{X}^{(t)}$ is:

Figure 1. Decomposing the covariance matrix Σ_p into Σ_p^- and $\vec{\rho}$ on a second-order neighborhood system ($K = 8$). By expressing both Φ_β and Ψ_β in terms of Kronocker products, it is possible to compute Fisher information in a efficient way during computational simulations.



$$\Psi_\beta = \frac{1}{\sigma^2} \|\Sigma_p^-\|_+ \quad (11)$$

Finally, using this matrix-vector notation, the maximum pseudo-likelihood estimator of the inverse temperature parameter can be rewritten as:

$$\hat{\beta}_{MPL} = \frac{\|\rho\|_+}{\|\Sigma_p^-\|_+} \quad (12)$$

2.6. Entropy in the GMRF Model

Our definition of entropy in the Gaussian Markov random field is done by repeating the same process employed to derive Φ_β and Ψ_β . Knowing that the entropy of random variable x is defined by the expected value of self-information, given by $-\log p(x)$, we have the following expression:

$$H_\beta = -E \left[\log p \left(x_i | \eta_i, \vec{\theta} \right) \right] = \frac{1}{2} \left[\log (2\pi\sigma^2) + 1 \right] - \frac{1}{\sigma^2} \left[\beta \sum_{j \in \eta_i} \sigma_{ij} - \frac{\beta^2}{2} \sum_{j \in \eta_i} \sum_{k \in \eta_i} \sigma_{jk} \right] \quad (13)$$

Using the same matrix-vector notation introduced in the previous sections, we can further simplify the expression for H_β .

Definition 5. Let an isotropic pairwise GMRF be defined on a lattice $S = \{s_1, s_2, \dots, s_n\}$ with a neighborhood system η_i . Assuming that $\mathbf{X}^{(t)} = \{x_1^{(t)}, x_2^{(t)}, \dots, x_n^{(t)}\}$ denotes the global configuration of the system at time t , and $\vec{\rho}$ and Σ_p^- are defined according to Figure 1, the entropy H_β for $\mathbf{X}^{(t)}$ is:

$$H_\beta = H_G - \frac{1}{\sigma^2} \left[\beta \|\vec{\rho}\|_+ - \frac{\beta^2}{2} \|\Sigma_p^-\|_+ \right] = H_G - \left[\frac{\beta}{\sigma^2} \|\vec{\rho}\|_+ - \frac{\beta^2}{2} \Psi_\beta \right] \quad (14)$$

where $H_G = 0.5 [\log(2\pi\sigma^2) + 1]$ denotes the entropy of a Gaussian random variable with variance σ^2 and Ψ_β is the type-II expected Fisher information. For more details the author is referred to [11].

2.7. Uncertainty in the Inverse Temperature Estimation

In estimating the inverse temperature parameter of random fields via maximum pseudo-likelihood, a relevant question emerges: how to measure the uncertainty in the estimation of β ? Is it possible to quantify this uncertainty? We will see that both versions of Fisher information play a central role in answering this question.

It is known from the statistical inference literature that both maximum likelihood and maximum pseudo-likelihood estimators share an important property: asymptotic normality [17,18]. It is possible, therefore, to characterize their behavior in the limiting case by knowing the asymptotic variance. A limitation from maximum pseudo-likelihood approach is that there is no result proving that this method is asymptotically efficient (maximum likelihood estimators have been shown to be asymptotically efficient since in the limiting case their variance reaches the Cramer-Rao lower bound). It is known that the asymptotic covariance matrix of maximum pseudo-likelihood estimators is given by [19]:

$$C(\vec{\theta}) = H^{-1}(\vec{\theta}) J(\vec{\theta}) H^{-1}(\vec{\theta}) \quad (15)$$

with

$$H(\vec{\theta}) = E \left[\nabla^2 \log L \left(\vec{\theta}; \mathbf{X}^{(t)} \right) \right] \quad (16)$$

$$J(\vec{\theta}) = Var \left[\nabla \log L \left(\vec{\theta}; \mathbf{X}^{(t)} \right) \right] \quad (17)$$

where H and J denote, respectively, the Jacobian and Hessian matrices regarding the logarithm of the pseudo-likelihood function. Considering the single inverse temperature parameter, β , we have the following definition for the asymptotic variance of the maximum pseudo-likelihood estimator:

$$v_\beta = \frac{\text{Var}_\beta \left[\frac{\partial}{\partial \beta} \log L(\vec{\theta}; \mathbf{X}^{(t)}) \right]}{E_\beta^2 \left[\frac{\partial^2}{\partial \beta^2} \log L(\vec{\theta}; \mathbf{X}^{(t)}) \right]} = \frac{E_\beta \left[\left(\frac{\partial}{\partial \beta} \log L(\vec{\theta}; \mathbf{X}^{(t)}) \right)^2 \right] - E_\beta^2 \left[\frac{\partial}{\partial \beta} \log L(\vec{\theta}; \mathbf{X}^{(t)}) \right]}{\left(-E_\beta \left[\frac{\partial^2}{\partial \beta^2} \log L(\vec{\theta}; \mathbf{X}^{(t)}) \right] \right)^2} \quad (18)$$

However, the expected value of the first derivative of $\log L(\vec{\theta}; \mathbf{X}^{(t)})$ (score function) with relation to β is zero:

$$E \left[\frac{\partial}{\partial \beta} \log L(\vec{\theta}; \mathbf{X}^{(t)}) \right] = \frac{1}{\sigma^2} \sum_{i=1}^n \left\{ E[x_i - \mu] - \beta \sum_{j \in \eta_i} E[x_j - \mu] \right\} = 0 \quad (19)$$

and the second term in the numerator vanishes, leading us to the final expression for v_β as a function of both type-I and type-II Fisher information, Φ_β and Ψ_β :

$$v_\beta = \frac{\Phi_\beta}{\Psi_\beta^2} = \frac{\Phi_\beta + \Psi_\beta - \Psi_\beta}{\Psi_\beta^2} = \frac{1}{\Psi_\beta} + \frac{1}{\Psi_\beta^2} (\Phi_\beta - \Psi_\beta) \quad (20)$$

showing that in the information equilibrium condition ($\Phi_\beta = \Psi_\beta$) we have the traditional Cramer-Rao lower bound, given by the inverse of the Fisher information. This information equality condition holds for models in the exponential family of distributions under certain regularity conditions (the differentiation and integration operators are interchangeable).

Therefore, we can compute the asymptotic variance of the MPL estimator of the inverse temperature parameter in Gaussian Markov random fields. A simple interpretation of this equation indicates that the uncertainty in the estimation of the inverse temperature parameter is minimized when Ψ_β is maximized and Φ_β is minimized. Essentially, it means that, in average, the local likelihood functions should not be flat (there is a reduced number of candidates for β) and most local patterns must be aligned to the expected global behavior. In the following, we provide a definition for the asymptotic variance of the inverse temperature MPL estimator in the Gaussian Markov random field model (using the matrix-vector notation).

Definition 6. Let an isotropic pairwise GMRF be defined on a lattice $S = \{s_1, s_2, \dots, s_n\}$ with a neighborhood system η_i . Assuming that $\mathbf{X}^{(t)} = \{x_1^{(t)}, x_2^{(t)}, \dots, x_n^{(t)}\}$ denotes the global configuration of the system at time t , and $\vec{\rho}$ and Σ_p^- are defined the same way as described in Figure 1, the asymptotic variance of the maximum pseudo-likelihood estimator of the inverse temperature parameter β is given by (using the same matrix-vector notation):

$$v_\beta = \frac{\sigma^2}{\|\Sigma_p^-\|_+} + \frac{1}{\|\Sigma_p^-\|_+^2} \left[2 \|\vec{\rho} \otimes \vec{\rho}^T\|_+ - 6\beta \|\vec{\rho}^T \otimes \Sigma_p^-\|_+ + 3\beta^2 \|\Sigma_p^- \otimes \Sigma_p^-\|_+ \right] \quad (21)$$

2.8. The Fisher Curve of a System

By computing Φ_β , Ψ_β and H_β we have access to three important information-theoretic measures regarding a global configuration of the random field, $\mathbf{X}^{(t)}$. The motivation behind the Fisher curve is the development of a computational tool for the study and characterization of random fields. Basically, the Fisher curve of a system is the parametric curve embedded in this information-theoretic space obtained by varying the inverse temperature parameter β from an initial value β_I to a final value β_F . The resulting curve provides a geometrical interpretation about how the random field evolves from a lower entropy configuration A to a higher entropy configuration B (or vice-versa), since the Fisher information plays an important role in providing a natural metric to the Riemannian manifold of a statistical model [4,5]. We will call the path from a global system configuration A to a global system configuration B as the *Fisher curve* (from A to B) of the system, denoted by $\vec{F}_A^B(\beta)$. Instead of using the notion of time as parameter to build the curve \vec{F} , we parametrize \vec{F} by the inverse temperature parameter β . In geometrical terms, we are measuring the deformation in one component of the metric tensor of the stochastic model induced by a displacement in the inverse temperature parameter direction.

Definition 7. Let an isotropic pairwise GMRF model be defined on a lattice $S = \{s_1, s_2, \dots, s_n\}$ with a neighborhood system η_i and $\mathbf{X}^{(\beta_1)}, \mathbf{X}^{(\beta_2)}, \dots, \mathbf{X}^{(\beta_n)}$ be a sequence of outcomes (global configurations) produced by different values of β_i (inverse temperature parameters) for which $A = \beta_{MIN} = \beta_1 < \beta_2 < \dots < \beta_n = \beta_{MAX} = B$. The system's Fisher curve from A to B is defined as the function $\vec{F} : \mathcal{R} \rightarrow \mathcal{R}^3$ that maps each configuration $\mathbf{X}^{(\beta_i)}$ to a point $(\Phi_{\beta_i}, \Psi_{\beta_i}, H_{\beta_i})$ in the information space, that is:

$$\vec{F}_A^B(\beta) = (\Phi_\beta, \Psi_\beta, H_\beta) \quad \beta = A, \dots, B \quad (22)$$

where Φ_β , Ψ_β and H_β denote the type-I expected Fisher information, the type-II expected Fisher information and the entropy of the global configuration $\mathbf{X}^{(\beta)}$, defined by equations (10), (11) and (14), respectively.

We are especially interested in characterizing random fields by measuring and quantifying their behavior as the inverse temperature parameter deviates from zero, that is, when temperature leaves infinity. As mentioned before, the isotropic pairwise GMRF model belongs to the regular exponential family of distributions when the inverse temperature parameter is zero ($T = \infty$). In this case, it has been shown that the geometric structure, whose natural Riemannian metric is given by the Fisher information matrix (metric tensor), has constant negative curvature (hyperbolic geometry). Besides, Fisher information can be measured by two different but equivalent ways (information equality).

As the inverse temperature increases, the model starts to deviate from this known scenario, and the original Riemannian metric does not correctly represents the geometric structure anymore (since there is an additional parameter). The manifold which used to be 2D (surface) now slowly is transformed (deformed) to a different structure. In other words, as this extra dimension is gradually emerging (since β not null), the metric tensor is transformed (the original 2×2 Fisher information matrix becomes a 3×3 matrix). We believe that the intrinsic notion of time in the evolution of a random field composed by Gaussian variables (isotropic pairwise GMRF) is caused by the irreversibility of this deformation process. In this particular study, we are concerned only in measuring the Fisher information regarding

the inverse temperature parameter (a single component of the metric tensor). A further investigation concerns the derivation, simulation and analysis of the complete Fisher information matrix in Gaussian random fields (the complete metric tensor). We intend to focus at this problem in future works.

3. Results and Discussion

In this section, we present some experimental results using computational methods for simulating the evolution of random fields. All the simulations were performed with Markov Chain Monte Carlo (MCMC) algorithms for generating random field outcomes based on the specification of the model parameters. In this paper, we make use of the Metropolis-Hastings algorithm [20].

Our main objective is to measure Φ_β , Ψ_β and H_β along a MCMC simulation in which the inverse temperature parameter β is controlled to guide the global system behavior. Initially, β is set to $\beta_{MIN} = 0$, that is, the initial temperature is infinite. In the following, β is linearly increased, with fixed increments $\Delta\beta$, up to an upper limit β_{MAX} . After that, the reverse process is performed, that is, the inverse temperature is linearly decreased using the same fixed increments ($-\Delta\beta$) down to zero. We are actually performing a positive displacement followed by a negative displacement along the inverse temperature parameter direction. By sensing a component of the metric tensor (Fisher information) at each point, we are trying to capture part of the deformation in the geometric structure of the manifold defined by the random field's parametric space throughout the process.

Our simulation was performed using the following parameter settings: $\mu = 0$, $\sigma^2 = 1$ (initial value), $A = \beta_{MIN} = 0$, $B = \beta_{MAX} = 0.5$, $\Delta\beta = 0.001$ and 1000 iterations. At each iteration, the values of μ and σ^2 are updated by computing the sample mean and sample variance, respectively. The inverse temperature parameter is updated by computing the maximum pseudo-likelihood estimative. Figures 2 and 3 show some samples of the random field during the evolution of the system, the real values of β (used to generate the random field outcomes) and the estimative $\hat{\beta}_{MPL}$ along the an entire MCMC simulation. Note that in this model, the maximum pseudo-likelihood estimator of the inverse temperature parameter underestimates the real parameter. The experimental results suggest that an upper bound for $\hat{\beta}_{MPL}$ is a value close to $1/K$, where K is the size of the neighborhood system. In all experiments we consider a second-order system, which corresponds to $K = 8$.

Figure 4 shows the asymptotic variance of the maximum pseudo-likelihood estimator of the inverse temperature parameter, given by equation (21), for the isotropic pairwise GMRF model along the MCMC simulation. Note that the critical issue concerning the bias-variance tradeoff in this model is the large bias (since the asymptotic variance is quite small in comparison to the bias).

We now proceed to the analysis of the entropy in the isotropic pairwise GMRF model. Figure 5 shows the behavior of H_β along the MCMC simulation. Note that entropy has some fluctuations for small values of β and it starts to increase as the inverse temperature parameter deviates from zero (or T deviates from infinity). In the beginning of the simulation, that is, for small values of β , H_β fluctuates around a base level. After a certain moment, the system's entropy shows a completely different behavior: it starts to increase rapidly. The simulation results suggest this behavior is directly related with variations in Fisher information (components of the metric tensor).

Figure 2. Global system configurations along a Markov chain Monte Carlo (MCMC) simulation. Evolution of the random field as the inverse temperature parameter, β , is changed in order to control the expected global behavior.

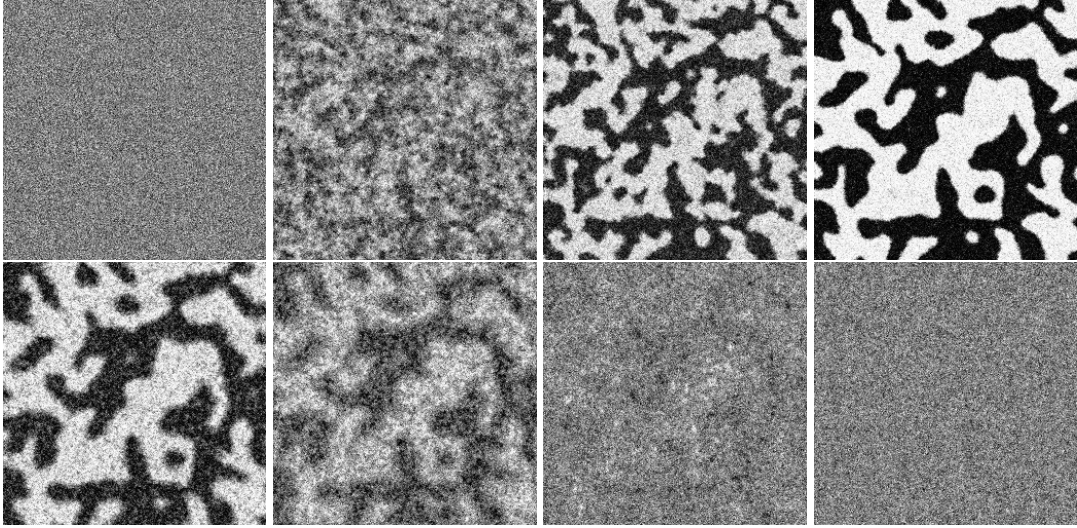


Figure 3. Variations in β and $\hat{\beta}_{MPL}$ along the MCMC simulation in the GMRF model. The MPL estimatives of the inverse temperature parameter in the isotropic pairwise GMRF model underestimates the real parameter values. The obtained results show that the upper bound for $\hat{\beta}_{MPL}$ is $1/K$, where K is the size of the neighborhood system.

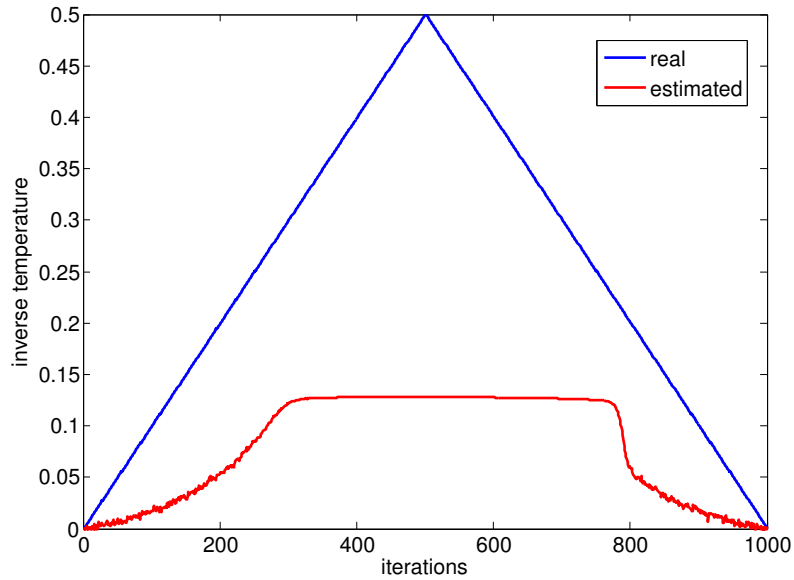


Figure 6 shows the plot of both forms of Fisher information, Φ_β and Ψ_β along the MCMC simulation. Some observations should be pointed out. First, it is clear that both forms of Fisher information significantly diverge for larger values of the inverse temperature parameter. In other words, the information equality prevails only when β is close to zero. Moreover, the asymmetry in both forms of Fisher information is clearly visible, even though the total displacement in the inverse temperature parameter direction is symmetric and adds up to zero. Note also that the way Φ_β and Ψ_β deviate from the equilibrium condition is significantly different from the way both of them approximate this condition.

Figure 4. Asymptotic variance of the maximum pseudo-likelihood estimator of the inverse temperature parameter along the MCMC simulation. In the isotropic pairwise GMRF model, the critical term in the bias-variance tradeoff is the large bias of $\hat{\beta}_{MPL}$, whose value is upper bounded by $1/K$, where K is the size of the neighborhood system.

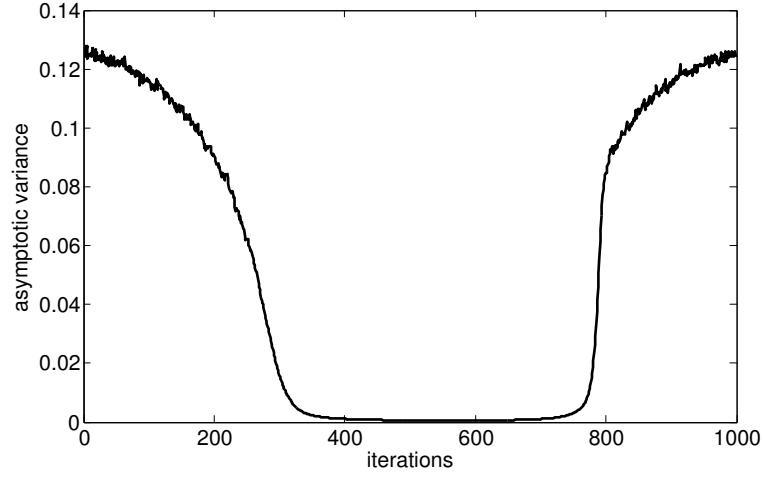
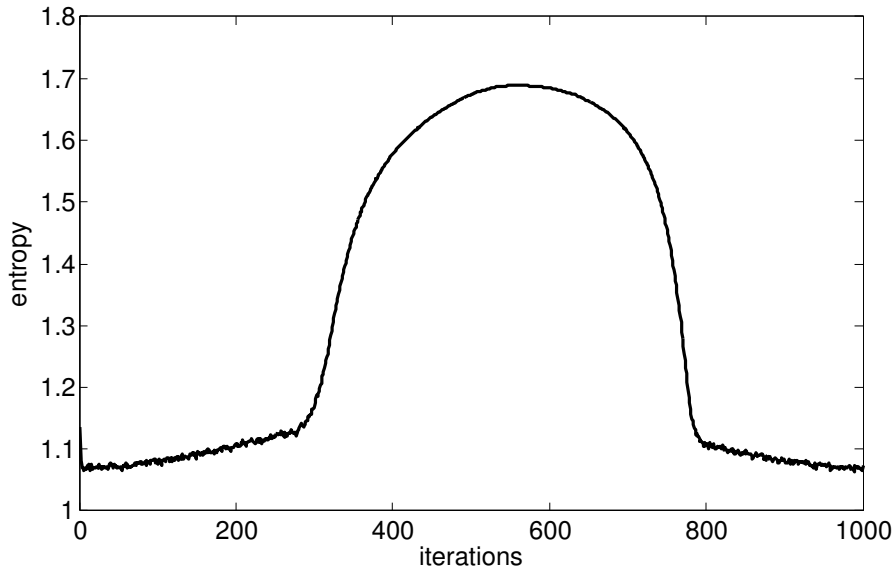
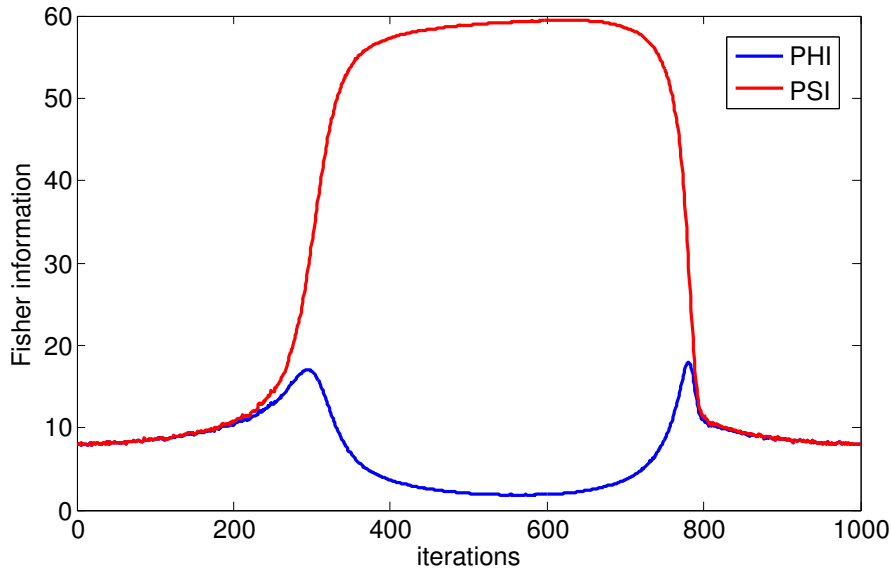


Figure 5. Entropy in the isotropic pairwise GMRF model along the MCMC simulation. When the temperature is infinity ($\beta = 0$), entropy reaches its minimum value. A close inspection to the graph reveals that the behavior of H_β is not symmetric, although the total displacement in the inverse temperature direction adds up to zero (from 0 to 0.5 and back).



In the following, a more geometric interpretation is discussed. In the beginning of the simulation, when the inverse temperature parameter is zero, the random field model degenerates to a simple Gaussian model (normal density). It is known that in this scenario the parametric space is a surface with constant curvature. In fact, this curvature is negative and equals minus one (hyperbolic geometry) [4]. The metric tensor, used to measure the deformation of the parametric space locally, is given by the Fisher information matrix. However, when temperature deviates from infinity (β deviates from zero), the original surface that represents the parametric space is transformed into a complex 3D Riemannian

Figure 6. Fisher information in the isotropic pairwise GMRF model along the MCMC simulation. When the temperature is infinity ($\beta = 0$), the information equality prevails, however, for larger values of β , Φ_β and Ψ_β diverge.

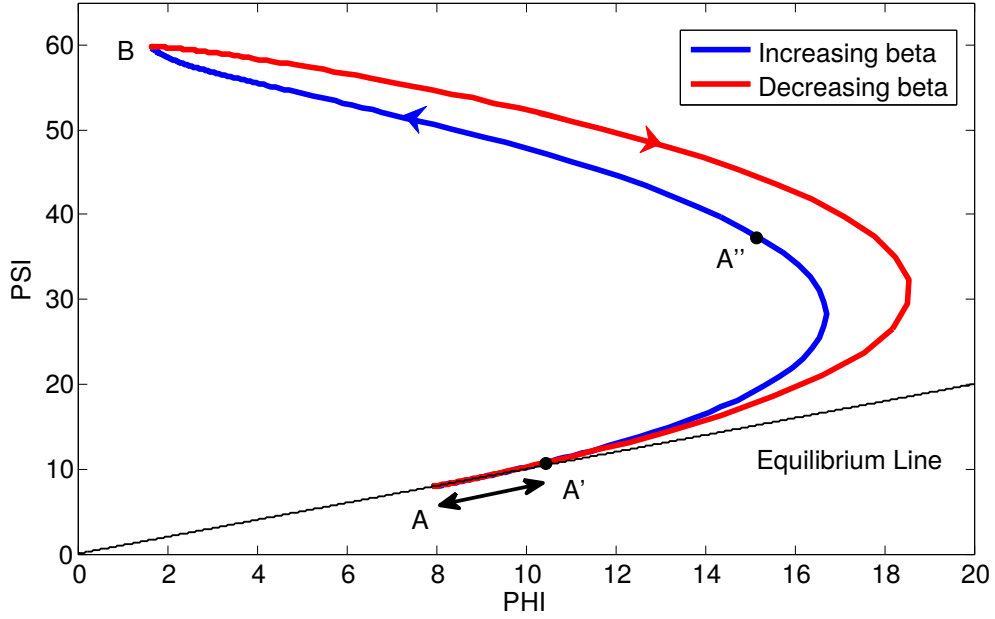


manifold, whose geometrical structure is defined by a novel metric tensor. The properties of this manifold are unknown, but by measuring Fisher information regarding the inverse temperature parameter we are trying to gain insights about the geometry of this manifold. In practical terms, what happens to the metric tensor can be summarized as: by moving forward δ units in β direction we sense an effect that is not compatible with the effect produced by a displacement of δ units in the opposite direction. In other words, moving towards higher entropy states (β increases) is different from moving towards lower entropy states (β decreases). This idea is illustrated by a plot of the Fisher curve of the random field along the simulation. Figure 7 shows the estimated 2D Fisher curves $\vec{F}_A^B(\beta) = (\Phi_\beta, \Psi_\beta)$ for $\beta = 0, \dots, 0.5$ (the blue curve) and $\vec{F}_B^A(\beta) = (\Phi_\beta, \Psi_\beta)$ for $\beta = 0.5, \dots, 0$ (the red curve).

Note that according to the Fisher curve, the basic notion of an arrow of time start to emerge when the information equilibrium condition vanishes. Suppose the following situation: initially, when temperature is infinite, the random field is at a state A. As temperature is gradually reduced (β is increasing), the system reaches a different state A'. During this period, there is no perception of time yet. Now, let us imagine that temperature is being increased (β is decreasing). In terms of information, the path from A' to A is the same as the path from A to A'. In other words, it is not possible to know whether we are moving forward or backwards in time, simply because at this point the notion of time is not clear (as well as the notions of past and future). Seemingly, time behaves as a space-like dimension since it is possible to move in both directions in this information space (the states A and A' are equivalent in terms of entropy, since there is no significant variation of H_β). Let us suppose now that from the state A the random field has evolved to the state A''. In this case, the notion of time is clear, since to take the system back to A it is necessary to go through a different path, that is, in terms of information the path from A to A'' is not the same as the path from A'' to A. In fact, at this point, the deformations induced by the metric tensor into the parametric space (manifold) are not reversible for opposite displacements in the inverse

temperature direction (it seems that the deformations caused in the manifold by the emergence of this extra dimension β are different from the deformations induced by the vanishing of this extra dimension).

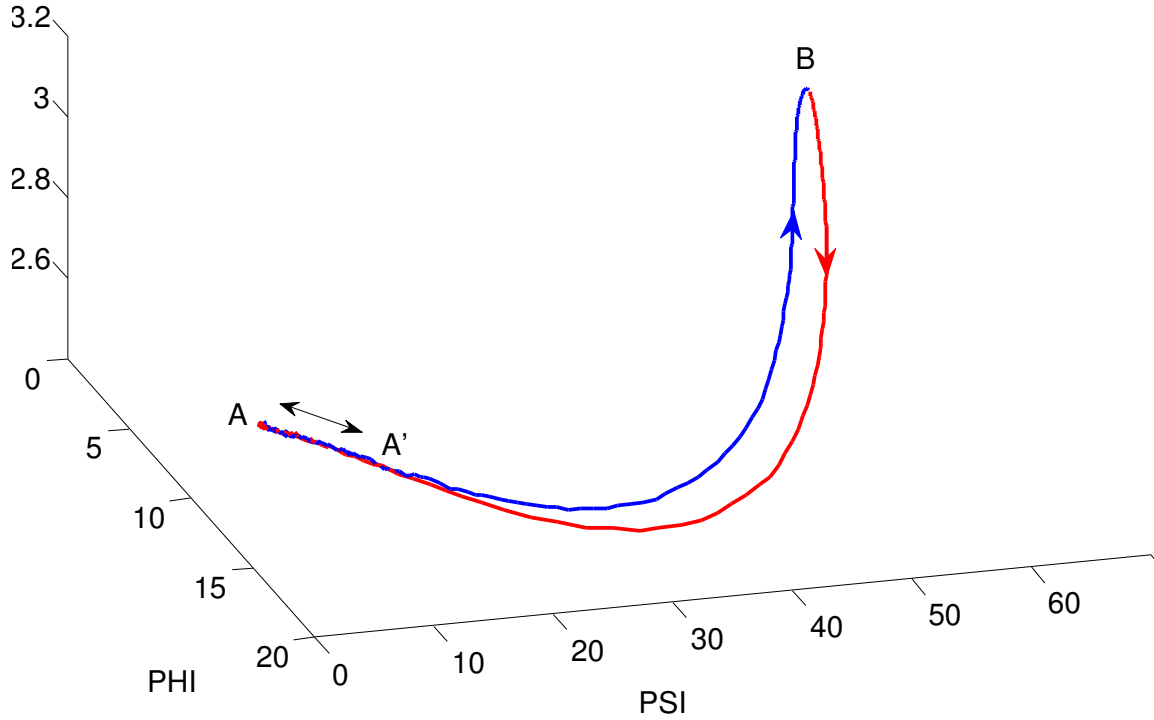
Figure 7. 2D Fisher curve of the random field along the MCMC simulation. The parametric curve was built by varying the inverse temperature parameter β from $\beta_{MIN} = 0$ to $\beta_{MAX} = 0.5$ and back. The results show that moving along different entropic states causes the emergence of a natural orientation in terms of information (the arrow of time).



The same curve illustrated by the previous Figure was plotted now in 3 dimensions. Figure 8 shows the 3D Fisher curve of the random field along the same MCMC simulation (now including entropy information). Once more, note that when the system is moving towards a higher entropy state (from A to B) the path is different from the one obtained by bringing the system back to a lower entropy state (from B to A). This natural orientation in the information space induces an arrow of time throughout the evolution of the random field. In other words, the only way to go from A to B by the red path would be running the simulation backwards. Note, however, that when moving along states whose variation in entropy is negligible (practically zero) the notion of time is not apparent. Again, suppose we move from the state A to the state A' indicated in Figure 8. The path from A to A' is the same as the path from A' to A, since both states are in the same base entropic level. In this case, it is not possible to detect whether we are moving forward or backwards in time (it is difficult to perceive the passage of time).

In order to emphasize the role of Fisher information in studying the emergence of the arrow of time in random fields, Figure 9 shows the parametric curve obtained by plotting some statistics used in the definition of both Φ_β and Ψ_β , more precisely, $\|\Sigma_p^-\|_+$ and $\|\vec{\rho}\|_+$. The obtained results show that by simply looking at these two measures it is not possible to capture an arrow of time. By analyzing these measurements we cannot say whether the system is moving forwards or backwards in time at all (even for large variations on the inverse temperature parameter). Note that the path from A ($\beta = 0$) to B ($\beta = 0.5$) is essentially the same as the path from B to A. Therefore, by monitoring only these two quantities we are not actually measuring the deformations induced by the metric tensor to the parametric space (random field model manifold).

Figure 8. 3D Fisher curve of the random field along the MCMC simulation. Note that, from a differential geometry perspective, as Φ_β deviates from Ψ_β the torsion of the curve becomes evident since it leaves the plane of constant entropy.



3.1. Final Remarks

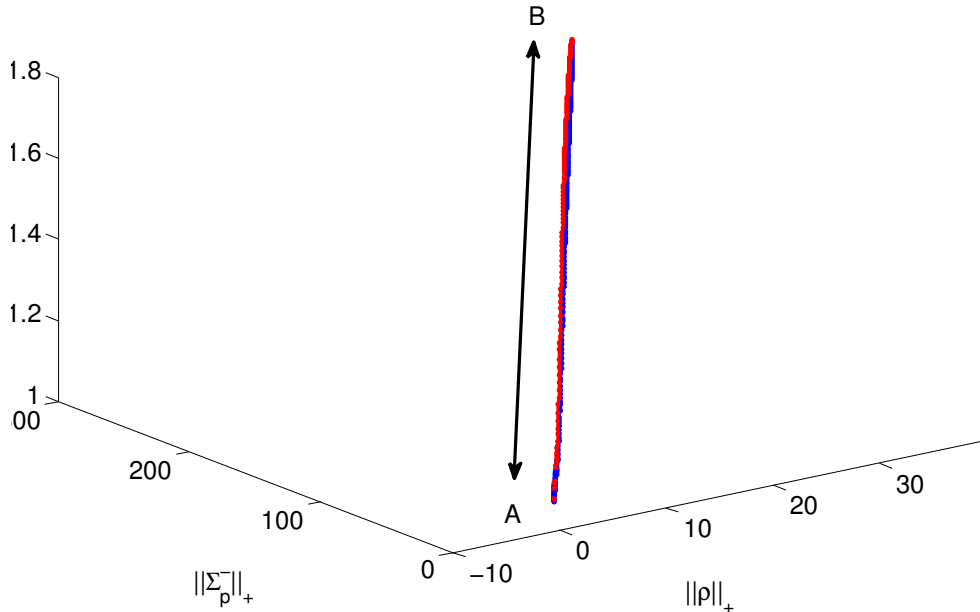
This section describes the main results obtained in this paper, focusing on the interpretation of the Fisher curve of a random field. Basically, when temperature is infinite entropy fluctuates around a minimum base value and the information equality prevails. From an information geometry perspective, a reduction in temperature (increase in β) causes a series of changes in the random field, since the metric tensor related to the parametric space is drastically deformed in an apparently non-reversible way, inducing the emergence of an arrow of time.

By quantifying and measuring an arrow of time in random fields, a relevant aspect that naturally arises concerns the notions of past and future. Suppose the random field is now in a state A, moving towards an increase in entropy (β is increasing). Within this context, the analysis of the Fisher curve suggests a possible interpretation: past is related to a set of states $P = \{X^{(\beta-)}\}$ whose entropies are lower than the entropy of the current state A. Or equivalently, past is also related to a set of states $P = \{X^{(\beta+)}\}$ whose entropies are higher than A, provided the random field is moving towards a lower entropy state.

Again, let us suppose the random field is in a state A and moving towards an increase in entropy (β is increasing). Similarly, future refers to a set of states $F = \{X^{(\beta+)}\}$ whose entropies are higher than the entropy of the current state A (or equivalently, future also refers to the set of states $F = \{X^{(\beta-)}\}$ whose entropies are lower than A, provided that the random field is moving towards a decrease in entropy).

Note that according to this possible interpretation, the notion of future is related to the direction of the movement, pointed by the tangent vector at a point $(\Phi_\beta, \Psi_\beta, H_\beta)$ of the Fisher curve. Therefore, if

Figure 9. Variations in the statistics $\|\Sigma_p^-\|_+$ and $\|\rho\|_+$ of an isotropic pairwise GMRF model along the MCMC simulation. The parametric curve was built by varying the inverse temperature parameter β from $\beta_{MIN} = 0$ to $\beta_{MAX} = 0.5$ and back. Note that in this case the arrow of time is not evident since the two curves, $\vec{F}_A^B(\beta)$ and $\vec{F}_B^A(\beta)$, are essentially the same.

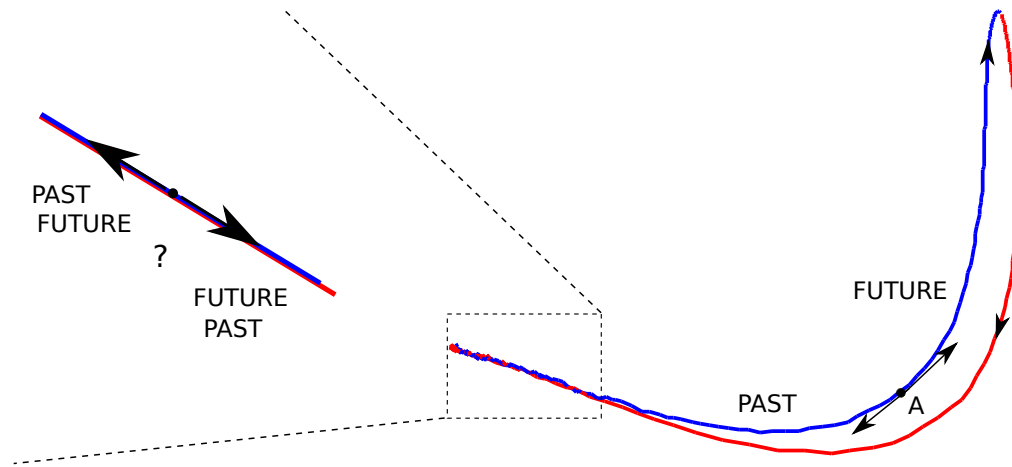


along the evolution of the random field there is no significant change in the system's entropy (as when the random field moves from A to A' in Figure 8), it would be possible to access past by simply moving into the opposite direction (by a displacement in the opposite direction along the inverse temperature parameter), as if time were a spatial dimension (given the stochastic nature of the model, this access to the past does not mean we would actually revisit exactly the same state again, only a equivalent one with the same global properties). However, in this case, where there is no significant changes in entropy, the notions of past and future seem to be meaningless.

4. Conclusions

In this paper, we addressed the problem of measuring the emergence of an arrow of time in Gaussian random field models. To intrinsically investigate the effect of the passage of time, we performed computational simulations of random fields in which the inverse temperature parameter is controlled to guide the system behavior throughout different entropic states. Investigations about the relation between two important information-theoretic measures, entropy and Fisher information, led us to the definition of the Fisher curve of a random field, a parametric trajectory embedded in an information space, which characterizes the system behavior in terms of variations in the inverse temperature parameter. Basically, this curve provides a geometrical tool for the analysis of random fields by showing how different entropic states are "linked" in terms of Fisher information, which is, by definition, the metric tensor of the underlying random field model parametric space. In other words, when the random field moves along different entropic states, its parametric space is actually being deformed by changes that happen in Fisher

Figure 10. The Fisher curve, an arrow of time and notions of past and future in the evolution of a Gaussian random field.



information (the metric tensor). In this scientific investigation we observe what happens to this geometric structure when the inverse temperature parameter is modified, that is, when temperature deviates from infinity, by measuring both entropy and Fisher information. An indirect subproblem involved in the solution of this main problem was the estimation of the inverse temperature parameter of a random field, given an outcome (snapshot) of the system. To tackle this subproblem, we used a statistical approach known as maximum pseudo-likelihood estimation, which is especially suitable for random fields, since it avoids computations with the joint Gibbs distribution, often computationally intractable. Our obtained results show that moving towards higher entropy states is different from moving towards lower entropy states, since the Fisher curves are not the same. This asymmetry induces a natural orientation to the process of taking the random field from an initial state A to a final state B and back, which is basically the direction pointed by the arrow of time, since the only way to move in the opposite direction is by running the simulations backwards.

Another observation regarding the analysis of the Fisher curves and the computational simulations suggest that the proposed intrinsic notion of time as the rate at which the parametric space is being deformed is highly non-linear. Apparently, the passage of time seems to be "faster" when entropy has an intermediate value (ΔH is large for small displacements in β) and "slower" when its value is extremal (very low or very high). In a Gaussian random field, it takes a greater effort to increase the system's entropy around a minimum base level in comparison to an intermediate value. The same observation is valid when decreasing the system's entropy: to move the system away from a maximum entropy state demands "more time" (ΔH is small for small displacements in β) since the deformations in the metric tensor are smooth and not abrupt (it is not so clear to perceive the passage of time).

Future works include the study of the Fisher curve in other random field models, such as the Ising and q-state Potts models. Moreover, we are currently working in the full specification of the Fisher information matrix regarding the Gaussian random field model (derivation of the all 9 components of the metric tensor) to completely characterize the geometrical structure of its manifold and to gain a deeper insight on the problems discussed here. Finally, a possible investigation concerns the analysis of the spectrum of the metric tensor throughout MCMC simulations. We expect that the eigenvalues of the

Fisher information matrix (a positive semidefinite operator) may reveal relevant information about the evolution of the system, since they will possibly be functions of the inverse temperature parameter.

Acknowledgements

The author would like to thank CNPQ (Brazilian Council for Research and Development) for the support through the research grant number 475054/2011-3.

Author Contributions

The main contribution of the paper is to investigate the Fisher curve as a geometrical tool to simulate Gaussian random field models whose temperature parameter deviates from infinity, allowing us to quantify the emergence of an arrow of time throughout the evolution of non-deterministic systems, by measuring deformations induced by the components of the metric tensor to the parametric space.

Conflicts of Interest

“The authors declare no conflict of interest”.

References

1. Haddad, W.M. Temporal Asymmetry, Entropic Irreversibility, and Finite-Time Thermodynamics: From Parmenides-Einstein Time-Reversal Symmetry to the Heraclitan Entropic Arrow of Time. *Entropy* **2012**, *14*, 407–455.
2. Jejjala, V.; Kavic, M.; Minic, D.; Tze, C.H. Modeling Time’s Arrow. *Entropy* **2012**, *14*, 614–629.
3. Campisi, M.; Hanggi, P. Fluctuation, Dissipation and the Arrow of Time. *Entropy* **2011**, *13*, 2024–2035.
4. Amari, S.; Nagaoka, H. *Methods of information geometry (Translations of mathematical monographs v. 191)*; American Mathematical Society, 2000.
5. Kass, R.E. The Geometry of Asymptotic Inference. *Statistical Science* **1989**, *4*, 188–234.
6. Lehmann, E.L. *Theory of Point Estimation*; Wiley: New York, 1983.
7. Bickel, P.J. *Mathematical Statistics*; Holden Day: New York, 1991.
8. Casella, G.; Berger, R.L. *Statistical Inference*, 2 ed.; Duxbury: New York, 2002.
9. Anandkumar, A.; Tong, L.; Swami, A. Detection of Gauss-Markov Random Fields with Nearest-Neighbor Dependency. *IEEE Trans. on Information Theory* **2009**, *55*, 816–827.
10. Gómez-Villegas, M.A.; Main, P.; Susi, R. The effect of block parameter perturbations in Gaussian Bayesian networks: Sensitivity and robustness. *Information Sciences* **2013**, *222*, 439 – 458. Including Special Section on New Trends in Ambient Intelligence and Bio-inspired Systems.
11. Levada, A.L.M. Learning from Complex Systems: On the Roles of Entropy and Fisher Information in Pairwise Isotropic Gaussian Markov Random Fields. *Entropy, Special Issue on Information Geometry* **2014**, *16*, 1002–1036.
12. Frieden, B.R. *Science from Fisher Information: A Unification*; Cambridge University Press: Cambridge, 2004.

13. Frieden, B.R.; Gatenby, R.A. *Exploratory Data Analysis Using Fisher Information*; Springer, 2006.
14. Hammersley, J.M.; Clifford, P. Markov field on finite graphs and lattices (unpublished).
15. Besag, J. Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society - Series B* **1974**, *36*, 192–236.
16. Isserlis, L. On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables. *Biometrika* **1918**, *12*, 134–139.
17. Jensen, J.; Kühnsh, H. On asymptotic normality of pseudo likelihood estimates for pairwise interaction processes. *Annals of the Institute of Statistical Mathematics* **1994**, *46*, 475–486.
18. Winkler, G. *Image Analysis, Random Fields and Markov Chain Monte Carlo Methods: A Mathematical Introduction*; Springer-Verlag New York, Inc.: Secaucus, NJ, USA, 2006.
19. Liang, G.; Yu, B. Maximum pseudo likelihood estimation in network tomography. *IEEE Trans. on Signal Processing* **2003**, *51*, 2043–2053.
20. Metropolis, N.; Rosenbluth, A.; Rosenbluth, M.; Teller, A.; Teller, E. Equation of state calculations by fast computing machines. *Journal of Chemical Physics* **1953**, *21*, 1087–1092.

© by the author; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).