Article

Gaussian Network's Dynamics Reflected into Geometric Entropy

Domenico Felice [1,2,*] **and Stefano Mancini** [1,2]

[1] School of Science and Technology, University of Camerino, I-62032 Camerino, Italy;
  E-Mail: stefano.mancini@unicam.it
[2] INFN-Sezione di Perugia, Via A. Pascoli, I-06123 Perugia, Italy

* Author to whom correspondence should be addressed; E-Mail: domenico.felice@unicam.it.

Abstract: We consider a geometric entropy as a measure of complexity for Gaussian networks, namely networks having Gaussian random variables sitting on vertices and their correlations as weighted links. We then show how the network dynamics described by the well-known Ornstein–Uhlenbeck process reflects into such a measure. We unveil a crossing of the entropy time behaviors between switching on and off links. Moreover, depending on the number of links switched on or off, the entropy time behavior can be non-monotonic.

Keywords: probability theory; Riemannian geometries; entropy; complex systems

1. Introduction

Many systems in the world can be modeled by a network, that is a set of items (vertices) with connections between them (links) [1]. Hence, the study of networks has been engaged in several branches of science. For example, in the mathematical form of graph theory, it is one of the fundamental grounds of discrete mathematics [2]. Moreover, networks have also been extensively studied in social science [3]. In recent years, network research moved from the analysis of single small graphs and the properties of individual vertices or links to the statistical properties of entire graphs [4]. As a consequence, several different statistical mechanics tools have been conceived of in order to quantify the different levels of organization of real networks [5]. Among them, entropic measures play a relevant role [6]. Along this line, by resorting to methods of information geometry [7], we have introduced recently a geometric entropy for Gaussian networks [8], namely networks having Gaussian random variables sitting

on vertices and their correlations as weighted links. Such network models, besides being rather easy to handle, find many applications, ranging from neural networks to wireless communication, from protein to electronic circuits, *etc.*

Traditionally, research on networks focused on features of static graphs. In this area, the study of topology networks is one of the most investigated issues. Although there is a wide diversity in the size and aim of networks observed in nature and technology, their topology shares several highly-reproducible and often universal characteristics [9]. Yet, almost all real networks are dynamic in nature [10]. Unfortunately, dealing with dynamic processes in network diversity prevails over the universality [11], and each network dynamical process is studied on its own terms, requiring its dedicated analytical formalism and numerical tools.

In the context of Gaussian networks, it seems quite natural to consider the dynamics of involved random variables described by a (multivariate) Ornstein–Uhlenbeck process [12]. In this way, the Gaussian character of the joint probability distribution is preserved over time. Taking this approach, we shall investigate how the evolution is reflected on the geometric entropy. Actually, we associate a Riemannian manifold endowed with a metric structure (that of Fisher–Rao) with the network at each time. Then, the change in level of organization of such an evolving system will be measured trough the change of the entropy, given by the logarithm of the volume of the manifold, when passing from one manifold to another.

As a simple case study, we shall consider a three-node Gaussian network. There, a first stochastic process will be considered with the aim of switching on all links starting from an uncorrelated Gaussian probability distribution (fully disconnected network). In contrast, a second stochastic process will be introduced aiming at switching off all links starting from a joint Gaussian probability distribution totally correlated (fully-connected network). The geometric entropy for these two cases varies monotonically as a function of time and shows a crossing behavior. Additionally, we shall deal with intermediate situations where part of the links are switched on and part switched off. In such cases, the geometric entropy no longer varies monotonically, while still exhibiting crossing behaviors between switched on and off links.

The layout of the article is as follows. In Section 2, we describe the measure of complexity related to the volume of the Riemannian statistical manifold associated with a static Gaussian network. In Section 3, we introduce the network dynamics, and we associate a manifold with the network at each time, thus providing a kind of foliation of the statistical manifold. Then, in Section 4, we study various dynamical cases occurring in a network having three vertices and show how their features are reflected into the geometric entropy. At the end, in Section 5, we draw our conclusions.

2. Statistical Models and Network Complexity Measure

We start considering n real-valued random variables $X_1, \ldots, X_n$ with joint Gaussian probability distribution:

$$p(x) = \frac{1}{\sqrt{(2\pi)^n \det C}} \exp\left[-\frac{1}{2}x^\top C^{-1} x\right] \tag{1}$$

where $x \in \mathbb{R}^n$ and C is the $n \times n$ covariance matrix with entries $C_{ij} = \mathbb{E}(X_i X_j)$, $i, j = 1, \ldots, n$. Throughout the paper, we assume to deal with zero mean probability distribution functions.

The set of all (zero mean) Gaussian probability distributions turns out to be a statistical model when we consider each function parametrized by its covariances $\mathbb{E}(X_i X_j)$, $i, j = 1, \ldots, n$. To be more precise, the function $p(x)$ in Equation (1) can be uniquely characterized by $m := \frac{n(n+1)}{2}$ parameters $\theta^l = \mathbb{E}(X_i X_j)$, with $l = \sum_{r=0}^{i-2}(n - r) + j - i + 1$ for $i \geq 2$ and $l = 1$ when $i = j = 1$. This amounts to number entries C_{ij} of the covariance matrix C when $j \geq i$. In such a way, we can rewrite $p(x) = p_\theta(x) \equiv p(x; \theta)$. Therefore, the set:

$$\mathcal{P} := \{p(x; \theta) \mid \theta \in \Theta \subset \mathbb{R}^m\} \tag{2}$$

is called the m-statistical multivariate (zero mean) Gaussian model on $\mathbb{R}^n$ as the mapping $\theta \mapsto p_\theta$ is injective, and the parameter space is $\Theta := \{\theta \in \mathbb{R}^m \mid C(\theta) > 0\}$ [7].

Given the statistical model $\mathcal{P}$ in Equation (2), the mapping $\varphi : \mathcal{P} \to \mathbb{R}^m$ defined by $\varphi(p_\theta) = \theta$ is injective, and it allows us to consider $\varphi = [\theta^l]$ as a coordinate system for $\mathcal{P}$. In addition, we assume that a change of coordinates $\psi : \Theta \to \psi(\Theta) \subset \mathbb{R}^m$ is such that the set $\{p_{\psi^{-1}(\kappa)} \mid \kappa \in \psi(\Theta)\}$, where κ is the set of new coordinates given by $\kappa^l := \psi(\theta^l)$, represents the same family of probability functions as $\mathcal{P} = \{p_\theta \mid \theta \in \Theta\}$. Finally, if we consider C^∞-differentiable changes of coordinates, then $\mathcal{P}$ can be considered as a C^∞ differentiable manifold, called a statistical manifold [7].

Consider now a point $\theta \in \Theta$; then, the Fisher information matrix of $\mathcal{P}$ at θ is the $m \times m$ matrix $g(\theta) = [g_{ij}(\theta)]$, where the entries are given by:

$$g_{ij}(\theta) := \int_{\mathbb{R}^n} dx \, p(x; \theta) \, \partial_i \ln p(x; \theta) \partial_j \ln p(x; \theta) \tag{3}$$

with ∂_i standing for $\frac{\partial}{\partial \theta^i}$ and i, j running in the set $\{1, \ldots, m\}$. The matrix $g(\theta)$ results in being symmetric and positive semidefinite. Yet, we assume from here on that $g(\theta)$ is positive definite. In such a way, we can endow the parameter space Θ with a Riemannian metric, the Fisher–Rao metric associated with the Gaussian statistical model $\mathcal{P}$ of Equation (1), given by $G(\theta) := \sum_{ij} g_{ij}(\theta) \, d\theta^i \otimes d\theta^j$, with g_{ij} as in Equation (3). Finally, the manifold $\mathcal{M} := (\Theta, G(\theta))$ describes a Riemannian manifold.

At this point, we can interpret the random variables $X_1, \ldots, X_n$ as sitting on vertices of a network and their correlations as weighted links among the vertices of such a network. Thereby, thanks to the parametric characterization of the Gaussian probability $p(x)$ in Equation (1), we can lift a network (discrete system) to a differentiable system (the Riemannian manifold $\mathcal{M} = (\Theta, G(\theta))$). To be more precise, we are considering random variables as hidden variables sitting on the vertices of an undirected network (without loops on its vertices), and their correlations are seen as weighted links among the vertices, again. Therefore, we are focusing the attention on the knowledge of some parameters characterizing the hidden variables. All of the information about the system is retained in these parameters. In particular, given the information on the variances and covariances of the multiple hidden variables, a multivariate Gaussian probability distribution describing the whole given network can be derived by means of the maximum entropy principle [13]. Thus, a parameter space is associated with any given network. This space encodes all of the information about the structure of the associated network.

In order to quantify the degree of organization of a given network, it is quite natural to consider the volume of the associated Riemannian manifold. To this end, let us first consider the volume element of the manifold $\mathcal{M} = (\Theta, G(\theta))$,

$$\nu_G = \sqrt{\det g(\theta)} \, d\theta^1 \wedge \ldots \wedge d\theta^m \tag{4}$$

Then, the volume $\mathcal{V}$ of $\mathcal{M}$ is evaluated as [8]:

$$\mathcal{V} := \int_{\Theta} \Upsilon(\theta)\, \nu_G \tag{5}$$

where $\Upsilon(\theta)$ is a suitable regularizing function introduced to avoid the possibility of having an ill-defined volume. Since the parameter space Θ is quite generally not compact, the function $\Upsilon(\theta)$ provides a kind of compactification of Θ by excluding the contributions of θ^l, making $\det g(\theta)$ divergent. It is clear that the regularizing function depends on the characteristic of the space Θ, as well as on the functional form of $\det g(\theta)$. An explicit expression of $\Upsilon(\theta)$ has been devised in some particular cases [8]. We are now in a position to define a complexity measure for a given network $\mathcal{X}$ with associated manifold $\mathcal{M} = (\Theta, G(\theta))$ as:

$$\mathcal{S} := -\ln \mathcal{V} \tag{6}$$

The definition (6) is inspired by the microcanonical definition of entropy in statistical mechanics [14]. In fact, the latter involves the phase space volume bounded by the hypersurface of constant energy E, which can be traced back to $\int_{M_E} \sqrt{\det g_J}\, d^n q$, where M_E is a configuration space subset determined by E, q are the configuration coordinates and g_J is the Jacobi kinetic energy tensor.

Unfortunately, computing such a quantity results in being quite hard. In fact, from Equations (1) and (3), it follows that:

$$\frac{1}{\sqrt{(2\pi)^n \det C}} \int dx\, f_{ij}(x) \exp\left[-\frac{1}{2} x^\top C^{-1} x\right] = \exp\left[\frac{1}{2} \sum_{\alpha,\beta=1}^{n} C_{\alpha\beta} \frac{\partial}{\partial x_\alpha} \frac{\partial}{\partial x_\beta}\right] f_{ij}|_{x=0} \tag{7}$$

where the exponential stands for a power series expansion over its argument (the differential operator) and:

$$f_{ij} := \partial_i \ln[p(x;\theta)]\, \partial_j \ln[p(x;\theta)] \tag{8}$$

The derivative of the logarithm reads:

$$\partial_i \ln[p(x;\theta)] = -\frac{1}{2}\left[\frac{\partial_i(\det C)}{\det C} + \sum_{\alpha,\beta=1}^{n} \partial_i\left((C^{-1})_{\alpha\beta}\right) x_\alpha x_\beta\right] \tag{9}$$

where $(C^{-1})_{\alpha\beta}$ denotes the $\alpha\beta$ entry of the inverse of the covariance matrix C in (1). The latter equation together with Equation (7) show the computational complexity of Equation (3). Indeed, the well-known formulas:

$$\partial_i C^{-1}(\theta) = C^{-1}(\theta)\big(\partial_i C(\theta)\big) C^{-1}(\theta) \tag{10}$$

$$\partial_i(\det C(\theta)) = \det C(\theta)\, \mathrm{Tr}(C(\theta)\, \partial_i(C(\theta))) \tag{11}$$

require the calculation of $m = \frac{n(n+1)}{2}$ derivatives with respect to the variables $\theta \in \Theta$ in order to work out the derivative of the logarithm in (9). Finally, to obtain the function f_{ij} in (8), we have to evaluate $O(n^4)$ derivatives. This quickly becomes an unfeasible task with growing n, even numerically.

To overcome this difficulty, we consider only the variances of the random variables as local coordinates of the manifold associated with the network. This means that the elements of the parameter space Θ are:

$$\theta \in \Theta \subset \mathbb{R}^m \Leftrightarrow \theta = (\theta^1,\ldots,\theta^m) \mid \theta^i = \mathbb{E}(X_i^2), \quad i \in \{1,\ldots,m\} \tag{12}$$

In this way, the topological dimension of the manifold $\mathcal{M} = (\Theta, G(\theta))$ reduces from $m = \frac{n(n+1)}{2}$ to $m = n$.

Thus, given a network $\mathcal{X}$ with n vertices and the $n \times n$ adjacency matrix A, consider the matrix $W = [W_{ij}]$ defined as:

$$W(\theta) = \text{diag}\left[\theta^1, \ldots, \theta^n\right] + R \tag{13}$$

with $\theta^i = \mathbb{E}(X_i^2)$. Here, $R = [R_{ij}]$ is the weighted $n \times n$ symmetric matrix associated with the network $\mathcal{X}$, whose entries are given by:

$$R_{ij} := \rho_{ij}\left(e_i^\top A e_j\right), \quad i, j \in \{1, \ldots, n\} \tag{14}$$

where ρ_{ij} is a real constant and e_i is the i-th element of the canonical basis in $\mathbb{R}^n$.

Finally, to the network $\mathcal{X}$, we associate the parameter space:

$$\widetilde{\Theta} := \{\theta \in \Theta \mid W(\theta) > 0\} \tag{15}$$

There, we introduce a new metric $\widetilde{G} = \sum_{ij} \widetilde{g}_{ij} d\theta^i \otimes d\theta^j$ with components:

$$\widetilde{g}_{ij} = \frac{1}{2}\left(\left(W^{-1}\right)_{ij}\right)^2 \tag{16}$$

where $\left(W^{-1}\right)_{ij}$ is the ij entry of the inverse of the matrix $W(\theta)$. This choice is motivated by the functional relation found in [8] between the Fisher–Rao matrix elements and the covariance matrix elements when the latter takes values of zero or one out of the main diagonal.

It turns out that $\widetilde{G}$ is a Riemannian metric. Hence, with a given network $\mathcal{X}$, we associate the Riemannian manifold $\widetilde{\mathcal{M}} = \left(\widetilde{\Theta}, \widetilde{G}(\theta)\right)$. Ultimately, we consider a statistical measure of the complexity of the network $\mathcal{X}$ with weighted matrix W and associated manifold $\widetilde{\mathcal{M}} = (\widetilde{\Theta}, \widetilde{G}(\theta))$ as:

$$\widetilde{\mathcal{S}}(W) := -\ln \widetilde{\mathcal{V}}(W) \tag{17}$$

where $\widetilde{\mathcal{V}}(W)$ is the volume of $\widetilde{\mathcal{M}}$ evaluated from the element:

$$\nu_{\widetilde{G}} := \sqrt{\det \widetilde{g}(\theta)}\, d\theta^1 \wedge \ldots \wedge d\theta^n \tag{18}$$

with $\widetilde{g}(\theta) = [\widetilde{g}_{ij}(\theta)]$ the $n \times n$ symmetric matrix whose entries are given by Equation (16). Notice, however, that also in this case, $\widetilde{\mathcal{V}}(W)$ might be ill-defined. The reason is two-fold: the set $\widetilde{\Theta}$ is not compact, because the variables θ^i are unbounded from above; moreover, $\det \widetilde{g}$ diverges, since $\det W$ approaches zero for some θ^i. Thus, we regularize it by introducing, as is usual in the literature [15], a regularizing function $\Upsilon(W(\theta))$ supplying a sort of compactification of the parameter space $\widetilde{\Theta}$ in order to avoid the possible divergences. We then have:

$$\widetilde{\mathcal{V}}(W) := \int_{\widetilde{\Theta}} \Upsilon(W(\theta))\, \nu_{\widetilde{G}} \tag{19}$$

3. Network Dynamics

Since we are dealing with Gaussian probability distribution functions and associated networks, it is quite natural to introduce a dynamics in terms of the Ornstein–Uhlenbeck process [12]. In fact, such a kind of process is known to maintain the Gaussian character of the distributions over time, and it has been already employed in neuronal networks (see, e.g., [16]). Additionally, a large variety of physical processes are described by such a stochastic evolution.

Recall that a stochastic process in $T \subset \mathbb{R}$ is a family $X = X(t)_{t \geq 0}$ of random variables. A multivariate Ornstein–Uhlenbeck process is a stochastic process defined by means of the following stochastic differential equation (SDE) [12]:

$$dX(t) = -\alpha(t)X(t)dt + \beta(t)dB(t) \tag{20}$$

where $\alpha(t)$ and $\beta(t)$ are $n \times n$ real square matrices usually referred to as drift and diffusion, respectively. Moreover, $dB(t)$ represents the infinitesimal stochastic Wiener increment. The solution of Equation (20) is the stochastic process $X(t)$ with mean:

$$\mathbb{E}(X(t)) = e^{-\int_0^t \alpha(t')dt'} \mathbb{E}(X(0)) \tag{21}$$

and covariance:

$$
\begin{aligned}
\mathbb{E}(X(t)X(t)^\top) &= e^{-\int_0^t \alpha(t')dt'} \mathbb{E}(X(0)X(0)^\top) \\
&\quad + \int_0^t dt' e^{-\int_{t'}^t \alpha(s)ds} \beta(t')\beta(t')^\top e^{-\int_{t'}^t \alpha(s)^\top ds}
\end{aligned} \tag{22}
$$

Here, $X(0)$ denotes the initial value of the SDE (20).

We are interested to the case of purely stochastic dynamics with no deterministic part, *i.e.*, $\alpha(t) \equiv 0$ for all $t \in [0, \infty)$, so that given $X(0)$ Gaussian distributed with zero mean, also $X(t)$ will result as such. Actually, by denoting with C_t the covariance matrix of $X(t)$, we obtain that the network is described at the time t by means of the following probability distribution function:

$$p(x; \theta_t) = \frac{1}{\sqrt{(2\pi)^n \det C_t}} \exp\left[-\frac{1}{2} x^\top C_t^{-1} x\right] \tag{23}$$

Here, the parameters $\theta_t = (\theta_t^1, \ldots, \theta_t^n)$ are exactly the variances of the random variables $X_1(t), \ldots X_n(t)$, *i.e.*, $\theta_t^i = \mathbb{E}(X_i(t)^2)$, evaluated at time t. They are the only parameters we are considering in order to characterize the probability density function in (23); in such a way, we can write $C_t \equiv C(\theta_t)$. The off-diagonal entries of the matrix C_t are understood as representing the weighted links among the vertices in which the random variables $X_i(t)$, $i = 1, \ldots, n$, sit. This amounts to describing the matrix C_t as the $n \times n$ weighted matrix W of Equation (13):

$$C_t \equiv W(\theta_t) = \mathrm{diag}\left[\theta_t^1, \ldots, \theta_t^n\right] + R(t) \tag{24}$$

where the entries $R_{ij}(t)$ of the matrix $R(t)$, previously defined in Equation (14), are given in this case by means of Equation (22) when $\alpha(t) \equiv 0$ and $i \neq j$,

$$R_{ij}(t) = \int_0^t dt' \left(\beta(t')\beta(t')^T\right)_{ij} \tag{25}$$

At this point, we can associate at each time $t \in [0, \infty)$ a manifold with the given network described by the probability density function (23). Hence, from Equations (15) and (16), with the network, at the time t, is associated the manifold $\widetilde{\mathcal{M}}_t = \left(\widetilde{\Theta}_t, \widetilde{G}(\theta_t) \right)$, where:

$$\widetilde{\Theta}_t := \left\{ \theta^t | C_t > 0 \right\} \tag{26}$$

and $\widetilde{G}(\theta_t) = \sum_{ij} \widetilde{g}_{ij}(\theta_t) \, d\theta_t^i \otimes d\theta_t^j$ with components given by:

$$\widetilde{g}_{ij}(\theta_t) = \frac{1}{2} \left(\left(W^{-1} \right)_{ij} \right)^2 \tag{27}$$

where $\left(W^{-1} \right)_{ij}$ is the ij entry of the inverse of the weighted matrix $W(\theta_t)$ in (24).

Finally, from Equation (17), we can quantify the level of organization of the dynamical network described by the SDE Equation (20) when $\alpha(t) \equiv 0$ by means of the following entropy:

$$\widetilde{\mathcal{S}}_t := \widetilde{\mathcal{S}}(C_t) = - \ln \widetilde{\mathcal{V}}(C_t), \quad t \in [0, \infty) \tag{28}$$

where $\widetilde{\mathcal{V}}(C_t)$ is the volume of the manifold $\widetilde{\mathcal{M}}_t$ associated with the network at the time t. It is worthwhile noticing that from Equation (24), we can write Equation (28) as depending on the covariance matrix C_t.

4. Trivariate Dynamical Networks: Results and Discussions

In order to get some insights from the model under study, we specialize to the statistical model $\mathcal{P}$ of Equation (2) with $m \equiv n = 3$. This entails a network with three vertices, so that the dynamics we are considering is described by a stochastic process $X(t) = (X_1(t), X_2(t), X_3(t)) \, (t \in [0, \infty))$ in $\mathbb{R}^3$ solution of (20) with $\alpha(t) \equiv 0$ for all $t \in [0, \infty)$.

Let us start considering an initial condition $X(0)$ characterized by the covariance matrix:

$$C_0 = \begin{pmatrix} \theta^1 & 0 & 0 \\ 0 & \theta^2 & 0 \\ 0 & 0 & \theta^3 \end{pmatrix} \tag{29}$$

That is, initially, we have a network with three vertices and no links between them. Then, let us consider a dynamics switching on all of the links, *i.e.*, the following diffusion matrix:

$$\beta(t) = \begin{pmatrix} 1 & -1 + \sqrt{1 + e^{-t}} & -1 + \sqrt{1 + e^{-t}} \\ -1 + \sqrt{1 + e^{-t}} & 1 & -1 + \sqrt{1 + e^{-t}} \\ -1 + \sqrt{1 + e^{-t}} & -1 + \sqrt{1 + e^{-t}} & 1 \end{pmatrix} \tag{30}$$

As a consequence of Equations (24) and (25), the Gaussian stochastic process $X(t)$ will be characterized by covariance matrix:

$$C_t = \begin{pmatrix} \theta_t^1 & 1 - e^{-t} & 1 - e^{-t} \\ 1 - e^{-t} & \theta_t^2 & 1 - e^{-t} \\ 1 - e^{-t} & 1 - e^{-t} & \theta_t^3 \end{pmatrix} \tag{31}$$

Let us remark that the latter covariance matrix C_t underlines a network with three vertices and the links among them having the same weights $R_{ij} = 1 - e^{-t}$ for $i \neq j$. The parameters θ_t^i's represent the

variances of the random variables $X_i(t)$ at the time t. They are the set of local coordinates that allow us to associate a Riemannian manifold with the network at the time t. As the time t goes to infinity, the network dynamics leads to a network with three vertices and all of the links with unit weight.

At this point, from Equation (26), we figure out the parameter space associated with the network at the time t. The conditions to have the matrix C_t positive definite are implemented by requiring that all of its principal minors are positive. Hence, in the case under consideration, we obtain:

$$\widetilde{\Theta}_t = \left\{ (\theta_t^1, \theta_t^2, \theta_t^3) \in \mathbb{R}^3 | \theta_t^1 > 0, \ \theta_t^2 > \frac{(1-e^{-t})^2}{\theta_t^1}, \ \theta_t^3 > (1-e^{-t})^2 \frac{\theta_t^1 + \theta_t^2 - 2(1-e^{-t})}{\theta_t^1 \theta_t^2 - (1-e^{-t})^2} \right\} \quad (32)$$

Subsequently, from Equation (27), we compute the metric:

$$
\begin{aligned}
\widetilde{G}(\theta_t) \ = \ & \frac{e^{2t}(-1+2e^t+e^{2t}(-1+\theta_t^2\theta_t^3))^2}{2(\det C_t)^2} d\theta_t^1 \otimes d\theta_t^1 + \frac{e^{2t}(-1+2e^t+e^{2t}(-1+\theta_t^1\theta_t^3))^2}{2(\det C_t)^2} d\theta_t^2 \otimes d\theta_t^2 \\
& + \frac{e^{2t}(-1+2e^t+e^{2t}(-1+\theta_t^1\theta_t^2))^2}{2(\det C_t)^2} d\theta_t^3 \otimes d\theta_t^3 + \frac{e^{2t}(-1+e^t)^2(1+e^t(-1+\theta_t^3))^2}{(\det C_t)^2} d\theta_t^1 \otimes d\theta_t^2 \\
& + \frac{e^{2t}(-1+e^t)^2(1+e^t(-1+\theta_t^2))^2}{(\det C_t)^2} d\theta_t^1 \otimes d\theta_t^3 + \frac{e^{2t}(-1+e^t)^2(1+e^t(-1+\theta_t^1))^2}{(\det C_t)^2} d\theta_t^2 \otimes d\theta_t^3
\end{aligned}
$$
$$(33)$$

Therefore, at the time t, the Riemannian manifold $\widetilde{\mathcal{M}}_t = (\widetilde{\Theta}_t, \widetilde{G}(\theta_t))$ turns out to be associated with the network.

We are now in a position to compute the volume of the manifold $\widetilde{\mathcal{M}}_t$ by applying Equation (19), where we set [8]:

$$\Upsilon(C_t) := e^{-\mathrm{Tr}(C_t)} \log \left[1 + (\det C_t)^2\right] \quad (34)$$

with $\mathrm{Tr}(C_t)$ denoting the trace of the matrix C_t. The choice of this regularizing function stems from the specific form of the metric $\widetilde{G}(\theta_t)$ in Equation (33). Indeed, in order to cure the divergence caused by the θ_t^i's making $\det C_t = 0$, we employed the logarithmic function; while we employed the exponential function to cure the divergence occurring as the θ_t^i's go to infinity.

Finally, letting t vary in $[0, \infty)$, we can describe the time-dependent behavior of the entropy $\widetilde{\mathcal{S}}_t$ of Equation (28). This is shown in Figure 1.

Let us now consider an initial condition $X(0)$ characterized by the covariance matrix:

$$C_0 = \begin{pmatrix} \theta^1 & 1 & 1 \\ 1 & \theta^2 & 1 \\ 1 & 1 & \theta^3 \end{pmatrix} \quad (35)$$

and a drift matrix:

$$\beta(t) = \begin{pmatrix} 1 & -1+\sqrt{1-e^{-t}} & -1+\sqrt{1-e^{-t}} \\ -1+\sqrt{1-e^{-t}} & 1 & -1+\sqrt{1-e^{-t}} \\ -1+\sqrt{1-e^{-t}} & -1+\sqrt{1-e^{-t}} & 1 \end{pmatrix} \quad (36)$$

The Gaussian stochastic process $X(t)$ is again worked out from Equations (24) and (25). It describes a dynamics starting from a fully-connected network and leading to a fully-disconnected network. The covariance matrix of $X(t)$ results:

$$C_t = \begin{pmatrix} \theta_t^1 & e^{-t} & e^{-t} \\ e^{-t} & \theta_t^2 & e^{-t} \\ e^{-t} & e^{-t} & \theta_t^3 \end{pmatrix} \tag{37}$$

By applying Equation (26), we figure out the parameter space associated with the network at the time t; it reads:

$$\widetilde{\Theta}_t = \left\{ (\theta_t^1, \theta_t^2, \theta_t^3) \in \mathbb{R}^3 | \theta_t^1 > 0, \ \theta_t^2 > \frac{e^{-2t}}{\theta_t^1}, \ \theta_t^3 > \frac{e^{-2t}(\theta_t^1 + \theta_t^2 - 2e^{-t})}{\theta_t^1 \theta_t^2 - e^{-2t}} \right\} \tag{38}$$

Then, from Equation (27), we get the metric:

$$\begin{aligned} \widetilde{G}(\theta_t) &= \frac{e^{2t}(-1 + e^{2t}\theta_t^2\theta_t^3)^2}{2(\det C_t)^2} d\theta_t^1 \otimes d\theta_t^1 + \frac{e^{2t}(-1 + e^{2t}\theta_t^1\theta_t^3)^2}{2(\det C_t)^2} d\theta_t^2 \otimes d\theta_t^2 \\ &+ \frac{e^{2t}(-1 + e^{2t}\theta_t^1\theta_t^2)^2}{2(\det C_t)^2} d\theta_t^3 \otimes d\theta_t^3 + \frac{(e^t - e^{2t}\theta_t^3)^2}{(\det C_t)^2} d\theta_t^1 \otimes d\theta_t^2 \\ &+ \frac{(e^t - e^{2t}\theta_t^2)^2}{(\det C_t)^2} d\theta_t^1 \otimes d\theta_t^3 + \frac{(e^t - e^{2t}\theta_t^1)^2}{(\det C_t)^2} d\theta_t^2 \otimes d\theta_t^3 \end{aligned} \tag{39}$$

The regularizing function of Equation (34) is a general function able to cure the divergences of the volume also in this case. Therefore, we compute the entropy $\widetilde{\mathcal{S}}_t$ in (28) by employing the function $\Upsilon(C_t)$ of Equation (34). The result is shown in Figure 1.
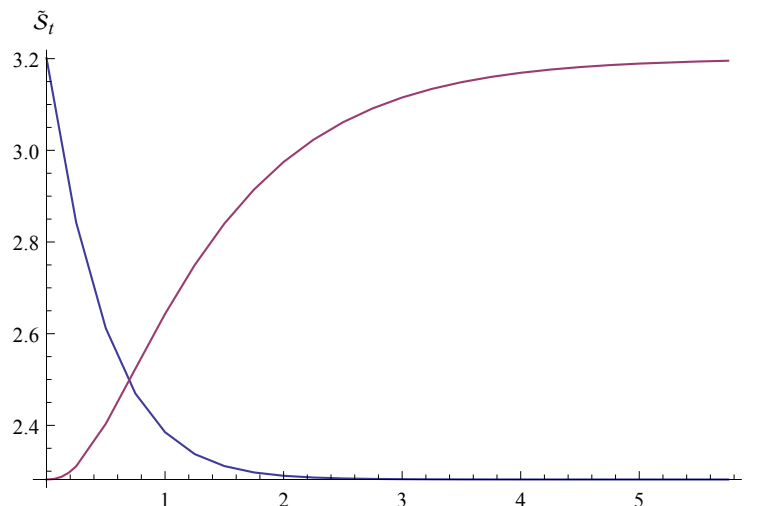

Figure 1. Geometric entropy $\widetilde{\mathcal{S}}_t$ of the dynamical network model when the initial state is a fully-disconnected network (magenta line) and when it is a fully-connected network (blue line).

In Figure 1, it is worth noting a crossing of the entropies' time behaviors related to the two cases (links switched on and off). The first dynamics we dealt with switches on the links between nodes and shows how the entropy monotonically increases *vs.* time. The maximum is asymptotically achieved and corresponds to the value computed for the fully-connected three-node network. The other dynamics

considered describes a fully-connected network that switches off all of the links. In this case, the entropy monotonically decreases *vs.* time. It asymptotically reaches the same value of that computed for the three-node networks with no links. Recall that the network dynamics is driven by the covariance matrices C_t given in (31) and (37). Then, at each time t, we consider a set of local coordinates $\{\theta_t^1, \theta_t^2, \theta_t^3\}$, and the time dependent entropy (28) is obtained as the logarithm of the volume of manifolds characterized by such a family (over t) of coordinates. The dynamics is reflected into this quantity in a very expected way; indeed, as is shown in [8], increasing or decreasing the topological dimension of the network, then the entropy increases or decreases, respectively. Therefore, adding links in the bare network amounts to growing the topological dimension up to the highest possible one represented by the fully-connected network. It is obvious that the contrary happens when the links are switched off from the fully-connected network.

Going further on, we consider intermediate situations with respect to the previous ones. That is, a network having initially one link and evolving in such a way that the other two links are switched on while the pre-existing one is switched off. This situation is characterized by, e.g.,

$$C_0 = \begin{pmatrix} \theta^1 & 1 & 0 \\ 1 & \theta^2 & 0 \\ 0 & 0 & \theta^3 \end{pmatrix} \tag{40}$$

and:

$$\beta(t) = \begin{pmatrix} 1 & -1 - \sqrt{1 - e^{-t}} & 1 + \sqrt{1 - e^{-t}} \\ -1 - \sqrt{1 - e^{-t}} & 1 & 1 + \sqrt{1 - e^{-t}} \\ 1 + \sqrt{1 - e^{-t}} & 1 + \sqrt{1 - e^{-t}} & 1 \end{pmatrix} \tag{41}$$

Furthermore, we consider a network having initially two links and evolving in such a way the other link is switched on while the pre-existing switched off. This situation is characterized by, e.g.,

$$C_0 = \begin{pmatrix} \theta^1 & 1 & 1 \\ 1 & \theta^2 & 0 \\ 1 & 0 & \theta^3 \end{pmatrix} \tag{42}$$

and:

$$\beta(t) = \begin{pmatrix} 1 & 1 + \sqrt{1 - e^{-t}} & 1 + \sqrt{1 - e^{-t}} \\ 1 + \sqrt{1 - e^{-t}} & 1 & -1 - \sqrt{1 - e^{-t}} \\ 1 + \sqrt{1 - e^{-t}} & -1 - \sqrt{1 - e^{-t}} & 1 \end{pmatrix} \tag{43}$$

In these cases, the entropy shows a non-monotonic time behavior, as can be seen in Figure 2. Actually, in both cases, $\widetilde{S}_t$ decreases towards a minimum and then rises up, reaching an asymptotic value. Each of these values corresponds to the initial value of the other curve. Here, the time-dependent entropy (28) reflects a more complex behavior than the one in Figure 1. Such a non-monotonic behavior could be explained by the fact that the dynamics described by Drifts (41), (43) involves two opposite processes, namely the switching on and off of links. Finally, it is worth highlighting that minimum values of the two curves are different; indeed, one is 2.4801 (magenta line), while the other is 2.47882 (red line).
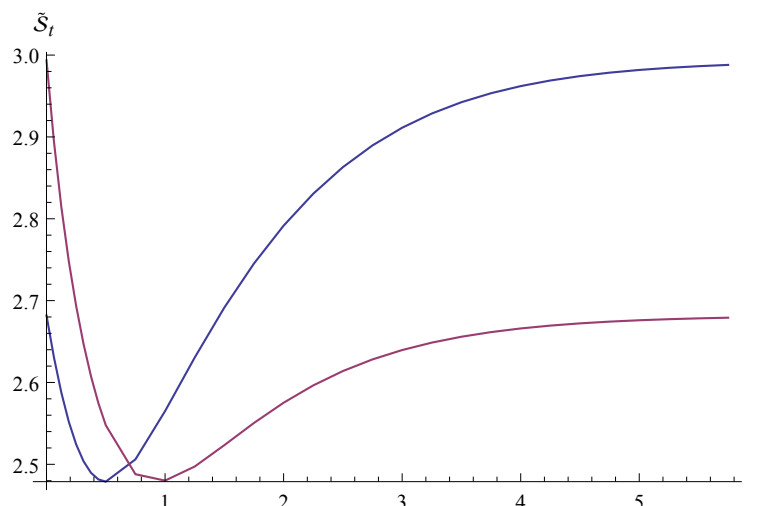
Figure 2. Geometric entropy $\widetilde{\mathcal{S}}_t$ of the dynamical network model when the initial state has two links (magenta line) and when it has one link (blue line).

5. Conclusions

In this work, we considered a way to associate a Riemannian manifold with a dynamical network interpreting random variables as sitting on vertices and their correlations as weighted links among vertices. We dealt with dynamics stemming from the Ornstein–Uhlenbeck process describing Gaussian networks' evolution in order to account for switching on and off links. This entails dealing with a family $X(t)$ of Gaussian random variables. In such a way, we associated a Riemannian manifold with the network at each time $t \in [0, \infty)$. It is worth remarking that at each time, a set of new local coordinates is provided. Hence, a mapping from $t \in [0, \infty)$ to $\mathbb{R}^+$ is supplied by the geometric entropy $\widetilde{\mathcal{S}}_t$ of Equation (28) in order to quantify the level of organization of the network along its time evolution.

Specifically, we considered a network with three vertices, and the dynamics is governed by Equation (20) setting the drift matrix $\alpha(t)$ to zero for every $t \in [0, \infty)$ and choosing the diffusion matrix $\beta(t)$ in order to describe the switching off and on of links.

The network's evolution is reflected in the time behavior of the geometric entropy in a simple monotonic way when all links are switched on or off together. In contrast, the time behavior of the geometric entropy becomes more involved (with minima) when part of the links is switched on and part off. In any case, a crossing of the entropy time behaviors between switching on and off links occurs.

It is worth noticing the difference with time-dependent entropy of [17] due to a network dynamics arising from an inference process [13]. In the present work, we started from a definition of static entropy [18] and then introduced a dynamics by the well-known Ornstein–Uhlenbeck process. In doing that, we considered a family of probability distributions and associate with it a network. Furthermore, the reverse process would be possible. In fact, given an n-vertexes network, we can always consider associated with it n random variables with correlations reflecting the edges disposition in the network and their weights. The problem is what kind of distribution do such variables have to respect. The choice could be done according to the type of network, or, better to say, to the kind of situation, or the system it represents. For instance, we may consider an electrical network where the edges are resistors and be interested in the voltage response on the current applied to each node. Then, voltages can be suitably

described by a Gaussian probability distribution, which accounts for thermal noise (Johnson–Nyquist noise).

A natural extension of the present work would contemplate the same dynamics in larger networks where the interplay of various links and their (de)-synchronization could lead to a more complex time dependence of $\widetilde{\mathcal{S}}_t$. Additionally, it would be interesting to go beyond purely stochastic dynamics and see how deterministic components affect the geometric entropy. This could be done, for example, by including deterministic pairwise interactions and eventually self-dynamics of node variables too, similar to those considered in [19].

Acknowledgments

Author Contributions

The authors have equally contributed to the paper. All authors have read and approved the final manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Newman, M.E.J. The structure and function of complex networks. *Siam Rev.* **2003**, *45*, 167–256.
2. Simonyi, G. Graph Entropy: A Survey. In *Combinatorial Optimization*; DIMACS Series in Discrete Mathematics and Theoretical Computer Science; Cook, W., Lovasz, L., Seymour, P., Eds.; American Mathematical Society: Providence, RI, USA, 1995; Volume 20, p. 399.
3. Borgatti, S.P.; Mehra, A.; Brass, D.J.; Labianca, G. Network Analysis in the Social Sciences. *Science* **2009**, *323*, 892–895.
4. Albert, R.; Barabasi, A.-L. Statistical Mechanics of Complex Networks. *Rev. Mod. Phys.* **2002**, *74*, 47–97.
5. Bianconi, G. Entropy of network ensembles. *Phys. Rev. E* **2009**, *79*, 036114.
6. Bianconi G. The entropy of randomized network ensembles. *Europhys. Lett.* **2008**, *81*, 28005.
7. Amari, S.; Nagaoka, H. *Methods of Information Geometry*; Oxford University Press: Oxford, UK, 2000.
8. Felice, D.; Mancini, S.; Pettini, M. Quantifying Networks Complexity from Information Geometry Viewpoint. *J. Math. Phys.* **2014**, *55*, 043505.
9. Strogatz, S.H. Exploring complex networks. *Nature* **2001**, *410*, 268–276.
10. Bilgin, C.C.; Yener, B. Dynamic Network Evolution: Models, Clustering, Anomaly Detection. Available online: https://www.cs.rpi.edu/research/pdf/08-08.pdf (accessed on 4 August 2015).

11. Dorogovtsev, S.N.; Goltsev, A.V. Critical phenomena in complex networks. *Rev. Mod. Phys.* **2008**, *80*, 1275–1335.

12. Da Prato, G.; Zabczyk, J. *Stochastic Equations in Infinite Dimensions*; Cambridge University Press: Cambridge, UK, 2008.

13. Cafaro, C. Works on an information geometrodynamical approach to chaos. *Chaos Solitons Fractals* **2009**, *41*, 886–891.

14. Pettini M. *Geometry and Topology in Hamiltonian Dynamics and Statistical Mechanics*; Springer: New York, NY, USA, 2007.

15. Leibbrandt, G. Introduction to the technique of dimensional regularization. *Rev. Mod. Phys.* **1975**, *47*, 849–876.

16. Grytskyy, D.; Tetzlaff, T.; Diesmann, M.; Helias, M. A unified view on weakly correlated recurrent networks. **2013**, arXiv:1304.7945.

17. Felice, D.; Cafaro, C.; Mancini, S. Information geometric complexity of a trivariate Gaussian statistical model. *Entropy* **2014**, *16*, 2944–2958.

18. Franzosi, R.; Felice, D.; Mancini, S.; Pettini, M. A geometric entropy measuring networks complexity. **2014**, arXiv:1410.5459.

19. Barzel, B.; Barabási, A.-L. Universality in networks dynamics. *Nature* **2013**, *9*, 673–681.