Article
Information Entropy Production of Maximum Entropy Markov Chains from Spike Trains

Rodrigo Cofre\textsuperscript{1} and Cesar Maldonado\textsuperscript{2}

\textsuperscript{1} CIMFAV, Facultad de Ingeniería, Universidad de Valparaíso, Valparaíso, Chile; rodrigo.cofre@uv.cl
\textsuperscript{2} IPICYT/División de Matemáticas Aplicadas, San Luis Potosí, Mexico; cesar.maldonado@ipicyt.edu.mx
* Correspondence: rodrigo.cofre@uv.cl

Abstract: Experimental recordings of the collective activity of interacting spiking neurons exhibit random behavior and memory effects, thus the stochastic process modeling the spiking activity is expected to show some degree of time irreversibility. We use the thermodynamic formalism to build a framework, in the context of spike train statistics, to quantify the degree of irreversibility of any parametric maximum entropy measure under arbitrary constraints, and provide an explicit formula for the information entropy production of the inferred Markov maximum entropy process. We provide examples to illustrate our results and discuss the importance of time irreversibility for modeling the spike train statistics.

Keywords: information entropy production; Discrete Markov Chains; spike train statistics; Gibbs measures; maximum entropy principle

1. Introduction

Since spike trains from experimental recordings are stochastic \cite{1}, and living systems are in non-equilibrium states (time irreversible) \cite{2}, a good candidate for a population coding scheme of living neuronal networks should be able to capture irreversibility in time \cite{3}. Thus, quantifying the degree of time irreversibility of spike trains becomes an important challenge which can be approach using tools from the fruitful intersection between information theory and statistical mechanics. Given a stochastic system, the quantity that measures how far it is from its equilibrium state (in statistical terms) is called information entropy production (IEP)\cite{4}.

In this paper, we quantify the IEP of parametric maximum entropy measures of populations of spiking neurons under arbitrary constraints. In Schneidman \textit{et al} \cite{5} and Pillow \textit{et al} \cite{6}, the authors used the maximum entropy principle focusing on firing rates and instantaneous pairwise interactions (Ising model) to describe the spike train statistics of the vertebrate retina responding to natural stimuli. Since then, the maximum entropy principle approach has become a standard tool to build probability measures in the field of spike train statistics \cite{5–8}. Recently, several extensions of the Ising model have been proposed, for example, the triplet model, considering as an extra constraint, the correlation of 3 neurons firing at the same time \cite{9}, and the so-called \textit{K}–pairwise model which consider \textit{K} neurons firing at the same time bin \cite{7} as an extra constraint. However, objections have appeared about their capability to predict time correlations. As discussed in \cite{10}, memory effects could have a non-negligible role in the spike train statistics. Nonetheless, most of the studies in this context have focused only on synchronous constraints and thus, modeling time-independent processes which are, by definition, reversible in time.

Since it is expected that memory effects show up in populations of spiking neurons, it is natural to ask about the information entropy production (IEP) associated with their statistical models. The maximum entropy approach can be extended to include non-synchronous constraints within the

\footnote{We distinguish the information entropy production with others forms of entropy production used in chemistry and physics.}
framework of the thermodynamic formalism and Gibbs measures in the sense of Bowen [11]. This opens the possibility to capture the irreversible character of the underlying biological process and thus, to provide statistical models biologically more realistic.

There is a vast body of theoretical work about irreversibility of stochastic processes, for mathematical details we refer the reader to [4]. In particular, for discrete time Markov chains, Gaspard [15] deduced an explicit expression for the change in entropy as the sum of a quantity called entropy flow plus the entropy production rate. In this paper, we follow this expression adapted to Markov chains associated with a finite range potential and we provide an explicit expression for the IEP of maximum entropy Markov chains.

This paper is organized as follows: In section 2 we introduce the setup of discrete homogeneous Markov chains, and review the properties that we use further. We present an example of the explicit computation of IEP in an integrate-and-fire spiking neuronal network model. In section 3 we introduce the maximum entropy principle within the framework of the thermodynamic formalism and Gibbs measures, discussing the role of the arbitrary constraints. We also provide the explicit formula to compute the IEP solely based on the spectral properties of the transfer matrix. In section 4 we provide examples of relevance in the context of spike train statistics. We finish this paper with discussions pointing out directions for further research.

2. Introduction

Since spike trains from experimental recordings are stochastic [1], and living systems are in non-equilibrium states (time irreversible) [2], a good candidate for a population coding scheme of living neuronal networks should be able to capture irreversibility in time [3]. Thus, quantifying the degree of time irreversibility of spike trains becomes an important challenge which can be approach using tools from the fruitful intersection between information theory and statistical mechanics. Given a stochastic system, the quantity that measures how far it is from its equilibrium state (in statistical terms) is called information entropy production (IEP) [4].

In this paper, we quantify the IEP of parametric maximum entropy measures of populations of spiking neurons under arbitrary constraints. In Schneidman et al [5] and Pillow et al [6], the authors used the maximum entropy principle focusing on firing rates and instantaneous pairwise interactions (Ising model) to describe the spike train statistics of the vertebrate retina responding to natural stimuli. Since then, the maximum entropy principle approach has become a standard tool to build probability measures in the field of spike train statistics [5–8]. Recently, several extensions of the Ising model have been proposed, for example, the triplet model, considering as an extra constraint, the correlation of 3 neurons firing at the same time [9], and the so-called K—pairwise model which consider K neurons firing at the same time bin [7] as an extra constraint. However, objections have appeared about their capability to predict time correlations. As discussed in [10], memory effects could have a non-negligible role in the spike train statistics. Nonetheless, most of the studies in this context have focused only on synchronous constraints and thus, modeling time-independent processes which are, by definition, reversible in time.

Since it is expected that memory effects show up in populations of spiking neurons, it is natural to ask about the information entropy production (IEP) associated with their statistical models. The maximum entropy approach can be extended to include non-synchronous constraints within the framework of the thermodynamic formalism and Gibbs measures in the sense of Bowen [11].

---

2 The notion of the Gibbs measure extends also to processes with infinite memory [12], and have been used in the context of spike train statistics [13,14].
3 We distinguish the information entropy production with others forms of entropy production used in chemistry and physics.
4 The notion of the Gibbs measure extends also to processes with infinite memory [12], and have been used in the context of spike train statistics [13,14].
opens the possibility to capture the irreversible character of the underlying biological process and thus, to provide statistical models biologically more realistic.

There is a vast body of theoretical work about irreversibility of stochastic processes, for mathematical details we refer the reader to [4]. In particular, for discrete time Markov chains, Gaspard [15] deduced an explicit expression for the change in entropy as the sum of a quantity called entropy flow plus the entropy production rate. In this paper, we follow this expression adapted to Markov chains associated with a finite range potential and we provide an explicit expression for the IEP of maximum entropy Markov chains.

This paper is organized as follows: In section 2 we introduce the setup of discrete homogeneous Markov chains, and review the properties that we use further. We present an example of the explicit computation of IEP in an integrate-and-fire spiking neuronal network model. In section 3 we introduce the maximum entropy principle within the framework of the thermodynamic formalism and Gibbs measures, discussing the role of the arbitrary constraints. We also provide the explicit formula to compute the IEP solely based on the spectral properties of the transfer matrix. In section 4 we provide examples of relevance in the context of spike train statistics. We finish this paper with discussions pointing out directions for further research.

3. Generalities

To set a common ground for the analysis of the IEP of spike trains, here we introduce the notations, and provide the basic definitions used throughout the paper.

3.1. Notation

We consider a finite network of $N \geq 2$ neurons. Let us assume that there is a natural time discretization such that at every time step, each neuron emits at most one spike. We denote the spiking-state of each neuron $\sigma^n_k = 1$ whenever the $k$-th neuron emits a spike at time $n$, and $\sigma^n_k = 0$ otherwise. The spike-state of the entire network at time $n$ is denoted by $\sigma^n = [\sigma^n_k]_{k=1}^N$, which we call a spiking pattern. For $n_1 \leq n_2$, we denote by $\sigma^{n_1,n_2}$ to an ordered concatenation of spike patterns $\sigma^{n_1,n_2} = \sigma^{n_1,1} \cdots \sigma^{n_2-1,n_2}$, that we call spike block. We call the sample of $T$ spiking patterns a spike train, which is a spike block $\sigma^{0,T}$.

We consider also infinite sequences of spike patterns that we denote $\vec{\sigma}$. We denote the set of infinite binary sequences of $N$ neurons $\Sigma_N$.

Let $L > 0$ be an integer, we write $\Sigma_N^L = \{0,1\}^{N \times L}$ for the set of spike blocks of $N$ neurons and length $L$. This is the set of $N \times L$ blocks whose entries are 0’s and 1’s. We introduce a symbolic representation to describe the spike blocks. Consider a fixed $N$, then to each spike block $\sigma^{0,L-1}$ we associate a unique number $\ell \in \mathbb{N}$, called block index:

$$\ell = \sum_{k=1}^N \sum_{n=0}^{L-1} 2^n N+k-1 \sigma^n_k.$$

We adopt the following convention: neurons are arranged from bottom to top and time runs from left to right in the spike train. For fixed $N$ and $L$, $\sigma^{(\ell)}$ is the unique spike block corresponding to the index $\ell$.

---

5 There is a minimal amount of time called “refractory period” in which no two spikes can occur. When binning, one could go beyond the refractory period and two spikes may occur in the same time bin. In those cases the convention is to consider only one spike.
3.2. Discrete-time Markov chains and spike train statistics

Let $\Sigma_N^L$ be the state space of a discrete time Markov chain, and let us for the moment use the following notation $\sigma(n) := \sigma^{n+L-1}$, for the random blocks and analogously $\omega(n) := \omega^{n+L-1}$ for the states. Consider the process $\{\sigma(n) : n \geq 0\}$. If $\sigma(n) = \omega(n)$ we say that the process is in the state $\omega(n)$ at time $n$. The transition probabilities are given as follows,

$$P[\sigma(n) = \omega(n) | \sigma(n-1) = \omega(n-1), \ldots, \sigma(0) = \omega(0)] = P[\sigma(n) = \omega(n) | \sigma(n-1) = \omega(n-1)].$$

We assume that this Markov chain is homogeneous, that is, (2) is independent of $n$. Consider two spike blocks $\sigma^{0,L-1}, \tilde{\sigma}^{1,L} \in \Sigma_N^L$ of length $L \geq 2$. Then the transition $\sigma(0) \rightarrow \tilde{\sigma}(1)$ is allowed if they have the common sub-block $\sigma^{1,L-1} = \tilde{\sigma}^{1,L-1}$.

We consider Markov transition matrices $P : \Sigma_N^L \times \Sigma_N^L \rightarrow \mathbb{R}$, whose entries are given by:

$$P[\sigma(1) | \sigma(0)] = \begin{cases} P[\tilde{\sigma}(1) | \sigma(0)] > 0 & \text{if } \sigma(0) \rightarrow \tilde{\sigma}(1) \text{ is allowed} \\ 0 & \text{otherwise.} \end{cases}$$

3.3. Detailed balance equations

Consider a fix $N$ and $L$. From the Markov property and the definition of the homogeneous transition matrix, one has for an initial measure $\nu$, the following Markov measure $\mu(\nu,P)$

$$\mu(\sigma(0) = \omega(0), \sigma(1) = \omega(1), \ldots, \sigma(k) = \omega(k)) = \nu(\omega(0))P_{\omega(0),\omega(1)} \cdots P_{\omega(k-1),\omega(k)},$$

for all $k > 0$. Here again, we used the short-hand notation $\sigma(k) := \sigma^{k,L+k-1}$ and $\omega(k) := \omega^{k,L+k-1}$.

An invariant probability measure of a Markov transition matrix $P$ is a row vector $\pi$ such that

$$\pi P = \pi.$$ (5)

We recall that for ergodic Markov chains (irreducible, aperiodic and positive recurrent) the invariant measure is unique.

Let us now consider a more general setting including non-stationary Markov chains. Let $\nu^0$ be the distribution of blocks $\sigma^{(\ell)} \in \Sigma_N^L$ at time $n$, then one has that the probability evolves in time as follows,

$$\nu^{n+1}(\sigma^{(\ell)}) = \sum_{\sigma^{(\ell')} \in \Sigma_N^L} \nu^n(\sigma^{(\ell')})P_{\sigma^{(\ell')},\sigma^{(\ell)}}.$$ (6)

For every $\sigma^{(\ell)} \in \Sigma_N^L$ one may write the following relation

$$\nu^{n+1}(\sigma^{(\ell)}) - \nu^n(\sigma^{(\ell)}) = \sum_{\sigma^{(\ell')} \in \Sigma_N^L} [\nu^n(\sigma^{(\ell')})P_{\sigma^{(\ell')},\sigma^{(\ell)}} - \nu^n(\sigma^{(\ell)})P_{\sigma^{(\ell')},\sigma^{(\ell)}}].$$

This last equation is related to the conditions of reversibility of a Markov chain. When stationarity and ergodicity are assumed, the unique stationary measure of the Markov chain $\pi$ is said to satisfy detailed balance if:

$$\pi_{\ell} P_{\ell',\ell} = \pi_{\ell'} P_{\ell',\ell} \quad \forall \sigma^{(\ell)}, \sigma^{(\ell')} \in \Sigma_N^L.$$ (7)
If the detailed balance equations are satisfied, then the quantity inside the parenthesis in the right-hand side of (6) is zero.

3.4. Information Entropy rate and Information Entropy Production

A well established measure of the amount of uncertainty of a probability measure $\nu$ is the information entropy rate, which we denote by $S(\nu)$. In the case of independent sequences of spike patterns ($L = 1$), the entropy rate is given by:

$$S(\nu) = - \sum_{\sigma^{(l)} \in \Sigma_N^L} \nu(\sigma^{(l)}) \log \nu(\sigma^{(l)}). \quad (8)$$

In the setting of ergodic stationary Markov chains taking values in the state space $\Sigma_N^L$, $L \geq 2$ with transition matrix $P$ and unique invariant measure $\pi$, the information entropy rate associated to the Markov measure $\mu(\pi, P)$ is given by:

$$S(\mu) = - \sum_{\sigma^{(l)}, \sigma^{(l')} \in \Sigma_N^L} \pi_{\ell} P_{\ell,\ell'} \log P_{\ell,\ell'}, \quad L \geq 2, \quad (9)$$

which corresponds to the Kolmogorov-Sinai entropy (KSE) [16].

Here, we introduce the information entropy production as in [15]. For expository reasons, let us consider again the non-stationary situation. The information entropy of a probability measure $\nu$ in the state space $\Sigma_N^L$ at time $n$ be given by

$$S_n(\nu) = - \sum_{\sigma^{(l)} \in \Sigma_N^L} \nu^n(\sigma^{(l)}) \log \nu^n(\sigma^{(l)}).$$

The change of entropy rate over one time-step is defined as follows:

$$\Delta S_n := S_{n+1}(\nu) - S_n(\nu) = - \sum_{\sigma^{(l)} \in \Sigma_N^L} \nu^{n+1}(\sigma^{(l)}) \log \nu^{n+1}(\sigma^{(l)}) + \sum_{\sigma^{(l)} \in \Sigma_N^L} \nu^n(\sigma^{(l)}) \log \nu^n(\sigma^{(l)}).$$

Arranging terms, one has that the previous equation can be written as:

$$\Delta S_n = - \sum_{\sigma^{(l)}, \sigma^{(l')} \in \Sigma_N^L} \nu^n(\sigma^{(l')}) P_{\ell,\ell'} \log \frac{\nu^{n+1}(\sigma^{(l)} P_{\ell,\ell'})}{\nu^n(\sigma^{(l)} P_{\ell,\ell'})} + \frac{1}{2} \sum_{\sigma^{(l)}, \sigma^{(l')} \in \Sigma_N^L} \left[ \nu^n(\sigma^{(l')}) P_{\ell,\ell'} - \nu^n(\sigma^{(l)} P_{\ell,\ell'}) \right] \log \frac{\nu^n(\sigma^{(l')} P_{\ell,\ell'})}{\nu^n(\sigma^{(l)} P_{\ell,\ell'})}, \quad (10)$$

the first part on the r.h.s of this equation is called information entropy flow and the second information entropy production [15].

Observe that in the stationary state, one has that $\nu^n = \nu^{n+1} = \pi$, thus the change of entropy rate is zero, meaning that information entropy flow equal information entropy production, therefore is possible to attain a steady state of fixed maximum entropy, but having positive IEP. In this case we refer to non-equilibrium steady state (NESS).

Here, since we are interested in the Markov chains that arise from the maximum entropy principle, we focus on the stationary case. In this case the IEP of a Markov measure $\mu(\pi, P)$ is explicitly given by:

$$IEP(P, \pi) = \frac{1}{2} \sum_{\sigma^{(l)}, \sigma^{(l')} \in \Sigma_N^L} \left[ \pi_{\ell'} P_{\ell,\ell'} - \pi_{\ell} P_{\ell,\ell'} \right] \log \frac{\pi_{\ell'} P_{\ell,\ell'}}{\pi_{\ell} P_{\ell,\ell'}} \geq 0, \quad (11)$$
nevertheless, we stress the fact that one can obtain the information entropy production rate also in the non-stationary case.

3.5. Example: Discrete time spiking neuronal network model

The integrate-and-fire model is one of the most ubiquitous models to simulate and analyze the dynamics of spiking neuronal circuits. This model is the simplest dynamical model that captures the basic properties of neurons, including the temporal integration of noisy sub-threshold inputs and all-or-nothing spiking. At the level of networks postulates a set of equations describing the behavior of the interconnected neurons motivated by the microscopic picture of how the biological neuronal network is supposed to work.

There exist several different versions of this model. Here we present the discrete time integrate-and-fire model. The model definition follows the presentation given in [17]. Neurons are considered as points, without spatial extension nor biophysical structure (axon, soma, dendrites). The dynamical system defined here is only ruled by discrete time dynamical variables.

Denote by \( V(t) \) the membrane potential vector with entries \( V_i(t) \), whose dynamics is defined as follows. Fix a real variable \( \theta > 0 \) called firing threshold. For a fixed discrete time \( t \), we have two possibilities:

- Either \( V_i(t) < \theta \), for all \( k = 1, \ldots, N \). This corresponds to sub-threshold dynamics.
- Or, there exists a \( k \) such that, \( V_k(t) \geq \theta \). Corresponding to firing dynamics.

The under-threshold dynamics is given by the following equation:

\[
V(t+1) = F(V(t)) + \sigma B(t)
\]

where

\[
F_i(V(t)) = \gamma V_i(t)(1 - Z[V_i(t)]) + \alpha \sum_{j=1}^{N} W_{ij} Z[V_j(t)] + \beta I_i.
\]

The function \( Z[x] := \mathbb{1}_{x \geq \theta} \) is called the firing state of neuron \( x \), where \( \mathbb{1} \) is the indicator function. When \( Z[V_i(t)] = 1 \) one says that neuron \( i \) spike otherwise is silent. We extend the definition of \( Z \) to vectors:

\[
Z[V(t)] = \mathbb{1}_{V(t) \geq \theta}
\]

\( Z[V_i(t)] \) is the vector with components \( Z[V_i(t)], i = 1, \ldots, N \). The leak rate is denoted by \( \gamma \in [0, 1] \), and \( W_{ij} \) is called the synaptic weight from the neuron \( j \) to the neuron \( i \). The synaptic weight is said to be excitatory if \( W_{ij} > 0 \) or inhibitory if \( W_{ij} < 0 \). The components of the vector \( B(t) \) are independent normalized Gaussian random variables and \( \sigma B \) is the noise amplitude parameter. The parameters \( \alpha \) and \( \beta \) were introduced in order to control the intensity of the synaptic weights and the stimulus, respectively.

With this model one obtains a set of conditional probabilities of spike patterns given the network’s spike history, allowing a mechanistic and causal interpretation of the origin of correlations (see [17] for details). Here, we consider only one time-step dependence on the past, although in the general approach it is possible to consider infinite memory. The conditional probabilities (transition matrix) are given as follows:

\[
P[\sigma | \sigma'] = \prod_{i=1}^{N} \left[ \sigma_i \varphi \left( \frac{\theta - C_i(\alpha, \beta, \sigma')}{\sigma B} \right) + (1 - \sigma_i) \left( 1 - \varphi \left( \frac{\theta - C_i(\alpha, \beta, \sigma')}{\sigma B} \right) \right) \right],
\]

where,

\[
C_i(\alpha, \beta, \sigma') = \gamma \alpha \sum_{j=1}^{N} W_{ij} \sigma'_j + \beta I_i
\]

and

\[
\varphi(x) = \int_{x}^{\infty} e^{-u^2} du.
\]
The function $C$ takes into account the past and the external stimuli (see [17] for details). These conditional (transition) probabilities define an ergodic Markov chain, specified by the biophysical dynamics of the spiking network. From the transition probabilities (14) and its unique steady state, we compute the IEP of this model using (11) for different values of the parameters $\alpha$ and $\beta$ (see Figure 1).

Figure 1. Plot of the average value of IEP for 500 realizations of the synaptic weight matrix for fixed $\alpha$ and $\beta$ in each case. We fix the following values of the parameters: $N = 6, \gamma = 0.2, c_0 = 1, \theta = 1, I_i = 1 \forall i \in \{1, ..., 6\}$. The components of the synaptic weight matrix $W_{ij}$ were drawn at random from a normalized Gaussian distribution. We plot the average value of IEP for 500 realizations of the synaptic weight matrix for fixed $\alpha$ and $\beta$ in each case.

Figure 1 shows that for this model the IEP depends mostly on the intensity of the synaptic weights, while the stimulus intensity is playing a minor role. This is an indication that IEP (in the stationary case) is essentially a property of the spiking neuronal network. The IEP of this neuronal network model is zero only under very restricted and unrealistic biophysical circumstances: when all synaptic weights are identical in amplitude and with the same sign or when they are all zero, i.e. when neurons do not communicate among them. In the first case spikes play a symmetrical role with respect to time, which cancels out when computing the IEP. In the second case the associated stochastic process is time independent, thus reversible. Therefore, generically this biophysically plausible model of spiking neuronal networks, has positive IEP. This means that the spike dynamics of this model leads to an irreversible Markov process.

4. Theoretical Framework

In the example of the previous section, we assume known the transition probabilities i.e., the structure of synaptic connectivity, stimulus and all other parameters defining the spiking neuronal network. Unfortunately this is not always the case. Usually, one only have access to a limited amount of experimental spiking data, which is a sampling of a very small subset of the entire state space. This makes that often the empirical frequencies are bad estimations of the elements of the Markov transition matrix. Here, we present how to use a variational principle from the thermodynamic formalism [18] to obtain the unique irreversible ergodic Markov transition matrix and its invariant measure having maximum entropy among those consistent with the constraints provided by data. This approach solves the problem of the bad estimations mentioned above and enables us to compute the IEP of the inferred Markov process, which is our main goal.

4.1. Inference of the maximum entropy Markov process

The problem of estimating the Markov chain of maximum entropy constrained by the data is of general interest in information theory. Consists in solving a constrained maximization problem, from
which one builds a Markov chain. The first step is choosing (arbitrarily) a set of indicator functions
(also called monomials) and determine from the data the empirical average of these functions. This
fixes the constraints of the maximization problem. After that, one maximizes the information entropy
rate, which is a concave functional in the space of Lagrange multipliers associated to the constraints,
obtaining the unique Markov measure that better approximates the statistics among all probability
measures that match exactly the constraints [19]. Up to our knowledge, previous approaches ignore
how to deal with the inference of irreversible Markov processes in the maximum entropy context
[20,21].

4.2. Observables and Potentials

Let us consider the space of infinite binary sequences $\Sigma_N$. An observable is a function $f : \Sigma_N \to \mathbb{R}$. We say that an observable $f$ has range $R$ if it depends only on $R$ consecutive spike patterns, e.g.

$$f(\sigma) = f(\sigma^{0,R-1}) .$$

We consider here that observables do not depend explicitly on time (time-translated invariant observables), i.e., for any time-step $n$, $f(\sigma^{0,n+R-1}) = f(\sigma^{n,n+R-1})$ whenever $\sigma^{0,R-1} = \sigma^{n,n+R-1}$.

Examples of observables are products of the form:

$$f(\sigma^{0,T}) = \prod_{u=1}^{r} \sigma_{k_u}^{n_u},$$

where $k_u = 1 \ldots N$ (neuron index) and $n_u = 0 \ldots T$ (time index). These observables are called monomials
and take values in $\{0,1\}$. Typical choices of monomials are $\sigma_{k_1}^{n_1}$ which is 1 if neuron $k_1$ fires at time $n_1$ and 0 otherwise; $\sigma_{k_1}^{n_1} \sigma_{k_2}^{n_2}$ which is 1 if neuron $k_1$ fires at time $n_1$ and neuron $k_2$ fires at time $n_2$ and 0 otherwise. For $N$ neurons and time range $R$ there are $2^{NR}$ possible monomials. To alleviate notations, instead of labeling monomials by a list of pairs, as in (17), we label them by an integer index, $l$ (the index is defined in the same way as the block index (1), i.e. a monomial reads $m_l$.

A potential is an observable that can be written as a linear combination of monomials$^6$. A potential of range $R$ is written as follows:

$$H(\sigma^{(l)}) := \sum_{i=1}^{2^{NR}} h_i m_i(\sigma^{(l)}) \quad \sigma^{(l)} \in \Sigma_N^R,$$

where the coefficients $h_i$ real numbers. Some coefficients in this series may be zero. We assume
throughout this paper that $h_i < \infty$$^7$. One example of potential is the one considering as monomials the
firing rates $\sigma_i$ and the synchronous pairwise correlations $\sigma_i \sigma_j$.

$$H(\sigma^{(l)}) = \sum_{i=1}^{N} h_i \sigma_i + \frac{1}{2} \sum_{i,j=1}^{N} l_{ij} \sigma_i \sigma_j \quad \sigma^{(l)} \in \Sigma_N^1$$

4.2.1. Additive observables of spike trains

Let $\phi$ be the shift map $\phi : \Sigma_N \to \Sigma_N$, defined by $\phi(\sigma)_{(i)} = \sigma_{(i+1)}$. Let $f$ be an arbitrary observable.

We may consider the sequence $\{ f \circ \phi (\sigma) \}$ as a random variable whose statistical properties depend
on those of the process producing the samples of $\sigma$ and the regularity of the observable $f$.

Given a spike train, one would like to empirically quantify properties empirical averages and
their fluctuation properties as a function of the sampling size. Consider a spike train $\sigma$, and let $n$ be the
sample length. The average of the observable $f$ of range $R \geq 1$ in $\sigma$ is given by,

---

$^6$ The range of the potential is the maximum over the ranges of the $m_l$ monomials considered.

$^7$ Here we do not consider hard core potentials with forbidden configurations.
\[ A_n(f) = \frac{1}{n - R + 1} \sum_{i=0}^{n-R} f \circ \phi^i(x), \]

in particular, for observables of range 1, one has

\[ A_n(f) = \frac{1}{n} \sum_{i=0}^{n-1} f(\sigma^i). \tag{19} \]

### 4.3. Variational principle

Let \( A_n(f_k) = C_k \) be the average value of \( K \) observables for \( k \in \{1, \ldots, K\} \). As the empirical average of monomials is not enough to uniquely determine the spike train statistics (there are infinitely many probability measures sharing the same averages of monomials), we use the maximum entropy method to obtain the Markov measure \( \mu \) that maximizes the KSE among all measures \( \nu \) that match the expected values of all observables, i.e. \( \nu[f_k] = C_k \), for all \( k \in \{1, \ldots, K\} \). This is equivalent to solve the following variational problem under constraints:

\[ S[\mu] = \max \left\{ S[\nu] : \nu[f_k] = C_k \ \forall \ k \in \{1, \ldots, K\} \right\}. \tag{20} \]

Since the function \( \nu \to S[\nu] \) is strictly concave, there is a unique maximizing Markov measure \( \mu(\pi, P) \) given the set of values \( C_k \). To solve this problem, we introduce the set of Lagrange multipliers \( h_k \in \mathbb{R} \) in the potential \( H = \sum_{k=1}^{K} h_k f_k \), which is a linear combination of the chosen observables. Next, we study the following unconstrained problem, which is a particular case of the so-called variational principle of the thermodynamic formalism [18]:

\[ \mathcal{P}[\mathcal{H}] = \sup_{\nu \in \mathcal{M}_{inv}} \left\{ S[\nu] + \nu[\mathcal{H}] \right\} = S[\mu] + \mu[\mathcal{H}], \tag{21} \]

where \( \mathcal{P}[\mathcal{H}] \) is called the free energy or topological pressure, \( \mathcal{M}_{inv} \) is the set of invariant measures with respect to the shift \( \phi \) and \( \nu[\mathcal{H}] = \sum_{k=1}^{K} h_k \nu[f_k] \) is the average value of \( \mathcal{H} \) with respect to \( \nu \).

In this paper, we only consider potentials \( \mathcal{H} \) of finite range, for which there is a unique measure \( \mu \) attaining the supremum [22] and is a **Gibbs measure in the sense of Bowen**.

#### Gibbs measures in the sense of Bowen.

Suppose \( \mathcal{H} \) is a finite range potential \( R \geq 2 \), a shift invariant probability measure \( \mu \) is called a Gibbs measure (in the sense of Bowen) if there are constants \( M > 1 \) and \( \mathcal{P}[\mathcal{H}] \in \mathbb{R} \) s.t.

\[ M^{-1} \leq \frac{\mu[\sigma^n]}{\exp(\sum_{k=1}^{K} h_k + R - 1) \mathcal{P}[\mathcal{H}] - (n + R - 1) \mathcal{P}[\mathcal{H}]}) \leq M \tag{22} \]

It is easy to see that the classical form of Boltzmann-Gibbs distributions \( \mu[\sigma] = e^{\mathcal{H}(\sigma)}/Z \) is a particular case of (22), when \( M = 1 \), \( \mathcal{H} \) is a potential of range \( R = 1 \) and \( \mathcal{P}[\mathcal{H}] = \log Z \).

### 4.3.1. Statistical Inference

The functional \( \mathcal{P}[\mathcal{H}] \) has the following property:

\[ \frac{\partial \mathcal{P}[\mathcal{H}]}{\partial h_k} = \mu[f_k] = C_k, \ \forall k \in \{1, \ldots, K\} \tag{23} \]

where \( \mu[f_k] \) is the the average of \( f_k \) with respect to \( \mu \), which is equal to the average value of \( f_k \) with respect to the empirical measure from the data \( C_k \), by constraint of the maximization problem. For finite range potentials \( \mathcal{P}[\mathcal{H}] \) is a convex function of \( h_l \)'s. This ensures the uniqueness of the solution of
with non-synchronous constraints [10].

4.4. Ruelle-Perron-Frobenius transfer operator

Consider $\mathcal{H}$ to be an arbitrary potential, and $w$ a continuous function on $\Sigma_N^r$. We introduce the Ruelle-Perron-Frobenius (R-P-F) transfer operator denoted by $\mathcal{L}_\mathcal{H}$, and it is given by,

$$\mathcal{L}_\mathcal{H} w(\sigma) = \sum_{\sigma' \in \Sigma_N^r \& \phi(\sigma') = \sigma} e^{\mathcal{H}(\sigma')} w(\sigma').$$

In an analogous way, as it is done for Markov approximations of Gibbs measures [23,24], for a finite range potential $\mathcal{H}$, we introduce the transfer matrix $\mathcal{L}_\mathcal{H}$,

$$\mathcal{L}_\mathcal{H}(\ell, \ell') = \begin{cases} e^{\mathcal{H}(\sigma^{0,\ell})} & \text{if } \sigma^{0,\ell} \sim \sigma^{(\ell)} \\ 0 & \text{otherwise.} \end{cases}$$

From the assumption $\mathcal{H} > -\infty$, each allowed transition corresponds to a positive entry in the matrix $\mathcal{L}_\mathcal{H}$.

4.5. Maximum entropy Markov chain for finite range potentials

The matrix (24) is primitive by construction, thus it satisfies the Perron-Frobenius theorem [25]. Let $\rho > 0$ be its spectral radius. Because of the irreducibility of the transfer matrix, $\rho$ is an eigenvalue of multiplicity 1 strictly larger in modulus than the other eigenvalues. For every $\sigma^{(\ell)} \in \Sigma_N^r$, let us denote by $L_\ell := L(\sigma^{(\ell)})$ and $R_\ell := R(\sigma^{(\ell)})$, the left and right eigenvectors of $\mathcal{L}_\mathcal{H}$ corresponding to the eigenvalue $\rho$. Notice that $L_\ell > 0$ and $R_\ell > 0$ for all $\sigma^{(\ell)} \in \Sigma_N^r$. Using spectral properties of the transfer matrix, we obtain the maximum entropy Markov transition probability matrix [22]:

$$P_{\ell,\ell'} := \frac{\mathcal{L}_\mathcal{H}(\ell, \ell') R_{\ell'}}{R_{\ell} \rho}, \quad \forall \sigma^{(\ell)}, \sigma^{(\ell')} \in \Sigma_N^r.$$  

The unique stationary probability measure $\pi$ associated to $P$ is also obtained by the spectral properties of $\mathcal{L}_\mathcal{H}$:

$$\pi_\ell := \frac{L_\ell R_\ell}{(L, R)}, \quad \forall \sigma^{(\ell)} \in \Sigma_N^r.$$  

For a finite range potential $\mathcal{H}$, the unique measure $\mu(\pi, P)$ associated to $\mathcal{H}$, satisfies the variational principle, furthermore, the topological pressure can be explicitly computed $P[\mathcal{H}] = \ln \rho$.

4.6. IEP of the inferred Markov maximum entropy process

Consider a potential $\mathcal{H}$ of finite range and the state space $\Sigma_N^r$. As we have seen before, using the maximum entropy framework one can build from the transfer matrix $\mathcal{L}_\mathcal{H}$, the Markov transition matrix $P$ and its invariant measure $\pi$. Furthermore, one can apply straightforwardly (25) and (26) to obtain a formula for the IEP based only on the spectral properties of $\mathcal{L}_\mathcal{H}$. After simplifying we obtain:

$$\text{IEP}(\mathcal{L}_\mathcal{H}) = \sum_{\sigma^{(\ell)}, \sigma^{(\ell')} \in \Sigma_N^r} \frac{L_\ell}{(L, R)} \frac{\mathcal{L}_\mathcal{H}(\ell, \ell') R_{\ell'}}{\rho} \log \left[ \frac{L_\ell R_{\ell'} \mathcal{L}_\mathcal{H}(\ell, \ell')}{L_{\ell'} R_{\ell} \mathcal{L}_\mathcal{H}(\ell', \ell)} \right]$$

8 The matrix $A$ is primitive if there is an $n \in \mathbb{N}$, s.t. $A^n$ has only positive components.
This is a quantity of major interest in spike train statistics, as it measure the degree of time irreversibility of the inferred entropy Markov chain. Although it is a straightforward result, it is quite general and of practical use, as we will see in the examples below. We can apply (25) and (26) to equation (7), we obtain the detailed balance condition in terms of the transfer matrix and its spectral properties:

\[ \frac{L_{\ell} R_{\ell'}}{\langle L, R \rangle} \frac{\mathcal{L}_H(\ell, \ell') R_{\ell'} \langle L, R \rangle}{L_{\ell'} R_{\ell'} \mathcal{L}_H(\ell', \ell)} = \frac{L_{\ell} R_{\ell'}}{\langle L, R \rangle} \frac{R_{\ell} \mathcal{L}_H(\ell', \ell)}{R_{\ell'} \mathcal{L}_H(\ell, \ell')} \]

Simplifying we obtain:

\[ \mathcal{L}_H(\ell, \ell') \mathcal{L}_H(\ell', \ell) = R_{\ell} \mathcal{L}_H(\ell, \ell') L_{\ell'} \]

4.7. Large deviations for observables of maximum entropy Markov chains

The Gärtner-Ellis theorem is very useful in our context, because it bypasses the direct calculation of the large deviation principle with rate function, for a given finite range observable \( f \) with respect to \( \mathbb{P} \). As we will see in the next section \( \lambda_f(k) \) naturally appears in the context of Maximum entropy Markov chains.

We define the \textit{scaled cumulant generating function} (SCGF) associated to the random variable \( f \) denoted by \( \lambda_f(k) \) as follows,

\[ \lambda_f(k) := \lim_{n \to \infty} \frac{1}{n} \ln \mathbb{E} \left[ e^{nkA_n(f)} \right], \quad k \in \mathbb{R}. \]

The \( n \)-th cumulant of the random variable \( f \) can be obtained by differentiating \( \lambda_f(k) \) with respect to \( k, n \) times and evaluating the result at \( k = 0 \). The next theorem by Gärtner-Ellis theorem relates the SCGF and the large deviations rate function. The Gärtner-Ellis theorem relies on the differentiability of \( \lambda_f(k) \), which is guaranteed for finite state Markov chains [26]. This theorem has several formulations, which usually require some technical definitions beforehand. Here we stated it in a simplified form, which is what we need for our purposes.

\textbf{Gärtner-Ellis theorem:} If \( \lambda_f(k) \) is differentiable, then there exist a large deviation principle for the average process \( A_n(f) \) whose rate function \( I_f(s) \) is the Legendre transform of \( \lambda_f(k) \):

\[ I_f(s) = \max_{k \in \mathbb{R}} \{ ks - \lambda_f(k) \} \]

The Gärtner-Ellis Theorem is very useful in our context, because it bypasses the direct calculation of \( \mathbb{P}(A_n(f)) \) in (29), i.e., having \( \lambda_f(k) \) a simple calculation leads to the rate function of \( f \). As we will see in the next section \( \lambda_f(k) \) naturally appears in the context of Maximum entropy Markov chains.

4.8. Large deviations for the IEP

Consider an irreducible Markov chain with transition matrix \( P_{\ell,\ell'} \). We define the \textit{tilted transition matrix by} \( f \) denoted by \( \hat{P}(f)(k) \), whose elements for a one time step observable are:

\[ \hat{P}(f)(k)_{\ell,\ell'}(k') = P_{\ell,\ell'} e^{k f(\ell')} \]
or for a two time step observable:

\[ \tilde{P}^{(f)}(\ell, \ell') = P_{\ell, \ell'} e^{kf_{\ell, \ell'}} \]  

For a Markov transition matrix \( P \) inferred from the maximum entropy, the tilted transition matrix can be built directly from the transfer matrix and its spectral properties.

\[ \tilde{P}^{(f)}(\ell, \ell') = \frac{L H(\ell, \ell') R_{\ell'} \rho}{R_{\ell} \rho} e^{kf_{\ell, \ell'}} \]  

The Markov chain structure underlying \( A_n(f) \) can be used here to obtain more explicit expressions for \( \lambda_f(k) \). In the case of the additive observables, if a Markov chain is homogeneous and ergodic can compute explicitly the SCGF as the logarithm of the maximum eigenvalue of \( \tilde{P}^{(f)} \):

\[ \lambda_f(k) = \ln(\rho(\tilde{P}^{(f)})) \]  

This result is valid if the state-space of the Markov chain is finite, where it can be proved furthermore that \( \lambda_f(k) \) is differentiable and \( \lambda_f(0) = \mathbb{E}(f) \).

Remark: The observable \( f \) does not need to belong in the set \( \{ f_k \}_{k=1}^K \) of chosen observables to fit the Markov maximum entropy process. We denote \( \rho(\tilde{P}^{(f)}) \) the dominant eigenvalue (i.e., with largest magnitude) of the matrix \( \tilde{P}^{(f)} \), which is unique by the Perron-Frobenius theorem.

We are interested in the fluctuations of the IEP. For that purpose we define the following observable:

\[ W_n(\{\sigma^i\}_{i=1}^n) = \ln \left[ \frac{P(\{\sigma^i\}_{i=1}^n)}{P(\{\sigma^i\}_{i=1}^{n-1})} \right] \]

where \( \{\sigma^i\}_{i=1}^{n-1} = \sigma^n, \sigma^{n-1}, \ldots, \sigma^1 \) is the temporal inversion of the trajectory \( \{\sigma^i\}_{i=1}^n \). It can be shown that for \( P \)-almost every trajectory of a stationary ergodic Markov chain \( (\pi, P) \):

\[ \lim_{n \to \infty} \frac{W_n(\{\sigma^i\}_{i=1}^n)}{n} = IEP(\pi, P) \]

It can be shown [4] that the SCGF \( \lambda_W(k) \) associated to the observable \( W_n \) can be found as the logarithm of the maximum eigenvalue \( \rho(k) \) of the matrix:

\[ \tilde{P}^{(W)}_{\ell, \ell'}(k) = P_{\ell, \ell'} e^{kF_{\ell, \ell'}} \]

where,

\[ F_{\ell, \ell'} = \ln \left[ \frac{\pi_{\ell} P_{\ell, \ell'}}{\pi_{\ell'} P_{\ell, \ell'}} \right] \]

which is a matrix of positive elements.

Using the Gärtner-Ellis theorem we obtain the rate function \( I_W(s) \) for the IEP observable:

\[ I_W(s) = \max_k \{ks - \lambda_W(k)\} \]

The rate function of the IEP observable has the following property:

\[ \lambda_W(k) = \lambda_W(-k - 1) \]
Since $\lambda_W'(0) = IEP(\pi, P)$ the symmetry implies

$$I_W(s) = I_W(-s) - s$$

4.8.1. Gallavotti-Cohen fluctuation theorem

The Gallavotti-Cohen fluctuation theorem refers to a symmetry in the fluctuations of the IEP. It is a statement about the large deviations of $W_\pi$, which is the time-averaged entropy production rate of the sample trajectory $\{\sigma_i\}_{i=1}^n$ of the Markov chain $\mu(\pi, P)$.

$$\frac{P[W_\pi \approx s]}{P[W_\pi \approx -s]} \approx e^{ns}$$

This means that the positive fluctuations of $W_\pi$ are exponentially more probable than negative fluctuations of equal magnitude. This is a universal ratio, i.e., no free parameters are involved and is experimentally observable.

5. Examples

In this section, we provide examples of applications of our results in the context of spike train statistics. We detail the transfer matrix technique to compute the maximum entropy Markov transition matrix and the invariant measure from a finite range potential $H$. We also compute the IEP and the fluctuations.

5.1. First example: Toy model

Consider a range-2 potential with $N = 2$ neurons:

$$H(\sigma^{0,1}) = h_1 \sigma_1^0 \sigma_0^1.$$

The transfer matrix (24) associated to $H$ is in this case a $4 \times 4$ matrix:

$$L_H = \begin{pmatrix}
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & e^{h_1} & 1 & e^{h_1} \\
1 & e^{h_1} & 1 & e^{h_1}
\end{pmatrix}.$$  

As this matrix is primitive by construction, it satisfies the hypothesis of the Perron-Frobenius theorem. Its unique maximum eigenvalue is $\rho = e^{h_1} + 3$. The left and right eigenvectors associated to this largest eigenvalue are respectively:

$$L \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \frac{2}{1 + e^{h_1}}; \quad L \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} = 1; \quad L \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \frac{2}{1 + e^{h_1}}; \quad L \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = 1,

$$

$$R \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \frac{2}{1 + e^{h_1}}; \quad R \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} = \frac{2}{1 + e^{h_1}}; \quad R \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} = 1; \quad R \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} = 1.$$

From the spectral properties of $L_H$ we obtain the Markov transition matrix (25), which reads,

$$P_{\sigma^0,\sigma^1} = \frac{1}{\hat{\rho}} \begin{pmatrix}
1 & 1 & \frac{1 + e^{h_1}}{2} & \frac{1 + e^{h_1}}{2} \\
1 & 1 & \frac{1 + e^{h_1}}{2} & \frac{1 + e^{h_1}}{2} \\
\frac{2}{1 + e^{h_1}} & \frac{2 e^{h_1}}{1 + e^{h_1}} & 1 & e^{h_1} \\
\frac{2}{1 + e^{h_1}} & \frac{2 e^{h_1}}{1 + e^{h_1}} & 1 & e^{h_1}
\end{pmatrix},$$
The unique invariant measure of this irreducible Markov transition matrix is given by equation (26), and its entries are given by,

\[ \pi \left( \begin{array}{c} 0 \\ 0 \end{array} \right) = \frac{4}{\rho^2}, \quad \pi \left( \begin{array}{c} 0 \\ 1 \end{array} \right) = \frac{2(\rho - 2)}{\rho^2}, \quad \pi \left( \begin{array}{c} 1 \\ 0 \end{array} \right) = \frac{2(\rho - 2)}{\rho^2}, \quad \pi \left( \begin{array}{c} 1 \\ 1 \end{array} \right) = \frac{(\rho - 2)^2}{\rho^2}. \]

It is easy to check that \( \pi \) is invariant w.r.t. the transition matrix \( P \), that is \( \pi P = \pi \).

From this example, we can verify that \textit{generically the detailed balance condition is not satisfied}; for example:

\[ P \left( \begin{array}{c} 0 \\ 1 \end{array} \right | \left. \begin{array}{c} 1 \\ 0 \end{array} \right) \pi \left( \begin{array}{c} 1 \\ 0 \end{array} \right) \neq P \left( \begin{array}{c} 1 \\ 0 \end{array} \right | \left. \begin{array}{c} 0 \\ 1 \end{array} \right) \pi \left( \begin{array}{c} 0 \\ 1 \end{array} \right). \]

As we can see in figure 2, the maximum entropy measure for the unconstrained problem is attained at the uniform distribution as expected \((h_1 = 0, \text{eigenvalue } \rho = 4 \text{ assigning probability } \frac{1}{4} \text{ to each spike pattern})\).

Let us now consider a constrained version of this problem. Suppose we have a data set of length \( T \) and we measure the average value of the observable \( f = \sigma_1^1 e_2^0 \),

\[ A_T(f) = 0.1 \]

Given this restriction and using the equation (23), we obtain the following equation:

\[ \frac{\partial \log(e^{h_1+3})}{\partial h_1} = 0.1 \]

Solving we find \( h_1 = -1.09861 \). Therefore, among all the Markov chains that match exactly the restriction, the one that maximizes the information entropy is the one obtained by fixing \( h_1 \) at the found value. Is easy to check that the variational principle (21) is satisfied.

From the transition probability matrix \( P \) and the invariant measure \( \pi \), we compute the KSE (9) and the IEP (27) as a function of the parameter \( h_1 \) (see figure 2).

**Figure 2.** IEP and KSE as a function of \( h_1 \). This figure illustrates the unconstrained problem, the maximum entropy measure is attained when \( h_1 = 0 \) i.e., the uniform distribution, which is also the Prigogine distribution of minimal IEP. In this example, the detailed balance condition is only satisfied in the trivial case \( h_1 = 0 \).
15 of 20

Figure 3. Gallavotti-Cohen fluctuation theorem for the first example at the parameter value $h_1 = -1$. Left) We show the SCGF associated to $W, \lambda_W(k)$, the derivative at zero is the IEP of the maximum entropy Markov chain, which in this case is 0.0557. This value coincides with the minimum of the rate function at the right side of the image.

5.2. Second example: Memoryless potentials

Consider a finite and fix number of neurons $N$ and a potential of range 1. This case includes the Ising model [5], Triplets [9], $K$-pairwise [7] and all other memoryless potentials that has been used in the context of maximum entropy models of spike train statistics. It represent a limit case in the definition of the transfer matrix, where transitions between spike patterns $\sigma \rightarrow \sigma'; \sigma, \sigma' \in \Sigma^1_N$ are considered and all transitions are allowed. In this case, the potential does not “see” the past i.e. $L_\mathcal{H}(\sigma, \sigma') = e^{\mathcal{H}(\sigma')}$. The matrix $L_\mathcal{H}$ has a unique maximum eigenvalue:

$$\lambda = Z = \sum_{\sigma' \in \Sigma^1_N} e^{\mathcal{H}(\sigma')}$$

and the rest of eigenvalues are equal to 0. The left and right eigenvectors corresponding to $\rho$ are:

$$L(\sigma') = \frac{1}{Z}, \quad R(\sigma') = e^{\mathcal{H}(\sigma')}; \quad \forall \sigma' \in \Sigma^1_N.$$ 

Note that $\langle L, R \rangle = 1$. We have therefore:

$$P(\sigma' | \sigma) = P(\sigma') = \pi(\sigma') = \frac{e^{\mathcal{H}(\sigma')}}{Z}; \quad \forall \sigma, \sigma' \in \Sigma^1_N,$$

(36) 

In this case, the invariant measure $\pi$ has the classical Boltzmann-Gibbs form. The associated Markov chain has no memory: successive events are independent. This last remark reflects a central weakness of memory-less maximum entropy models to describe neuronal dynamics.

Taking the formula of IEP (27) we obtain:

$$IEP(L_\mathcal{H}) = \sum_{\sigma, \sigma' \in \Sigma^1_N} \frac{L(\sigma) e^{\mathcal{H}(\sigma')} R(\sigma')}{\langle L, R \rangle \log(Z)} \left( \mathcal{H}(\sigma') - \mathcal{H}(\sigma) \right) = 0.$$ 

In the case where only range 1 observables are chosen (firing rates, pairwise correlations, triplets, etc.), the average value of these observables in a given data set is the same as the one taken from another data set where the time indexes have been randomly shuffled or even time reversed. As this is the only information about the process that the maximum entropy principle consider, it is not surprising that the stochastic process associated with the maximum entropy measure is time reversible. Consider a
data set consisting in binary patterns $D^O$. Let $g: \{0, \ldots, T\} \to \{0, \ldots, T\}$ be a function that randomly shuffles the time indexes, we call $D^{RS}$ the data set obtained after this transformation. Finally consider $D^I$, the data set with inverted time indexes,

$$D^O = \{\sigma^0, \sigma^1, \sigma^2, \ldots, \sigma^{T-1}, \sigma^T\}$$

$$D^{RS} = \{\sigma^{g(0)}, \sigma^{g(1)}, \sigma^{g(2)}, \ldots, \sigma^{g(T-1)}, \sigma^{g(T)}\}$$

$$D^I = \{\sigma^T, \sigma^{T-1}, \sigma^{T-2}, \ldots, \sigma^1, \sigma^0\}.$$

Observe that in these three cases (that may correspond to very different biological experiments), the average value of every observable of range one is exactly the same, therefore these data sets are characterized by the same maximum entropy distribution.

5.3. Third example: 1-time step Markov

Here, we consider the 1-time step extension of the Ising model, that reads:

$$\mathcal{H}(\sigma^{0:1}) = \sum_{i=1}^{N} h_i \sigma_i + \frac{1}{2} \sum_{i,j=1}^{N} J_{ij} \sigma_i \sigma_j + \sum_{i,j=1}^{N} f_{ij} \sigma_i \sigma_j.$$  (37)

This is the potential considered to fit a maximum entropy distribution to spiking data from a mammalian parietal cortex in-vivo in [27]. It is important to notice that in [27], the authors compute the solution of the maximum entropy problem imposing detailed balance condition, so in their case, there is zero IEP by construction. Here we do not consider a particular data set, instead we investigate the capability of this potential to generate IEP by considering the following scenarios: We consider a network of $N = 10$ neurons, where we draw at random the coefficients $h_i$ and $J_{ij}$ in a range plausible to be the maximum entropy coefficients (or Lagrange multipliers) of an experiment of retinal ganglion cells exposed to natural stimuli (values of from $h_i$ and $J_{ij}$ as in [28]). We generate the matrix $\gamma_{ij}$ by drawing each component at random from Gaussian distributions with different means and standard deviations. We summarize our results in figure 3. We observe the following: Independent of $h_i$ and $J_{ij}$ and the parameters of mean and variance from which the matrix of coefficients $\gamma_{ij}$ is generated, if $\gamma_{ij}$ is symmetric the Markov process generated by the potential (37) is reversible in time so the IEP is zero. This includes the limit case when $\gamma_{ij} = 0$, $\forall i, j \in \{1, \ldots, N\}$, where we recover the Ising model. Next, we fix the values of $h_i$ and $J_{ij}$ (random values), and we generate 100 matrices $\gamma_{ij}$ by drawing their components from Gaussian distributions $\mathcal{N}(0, \sigma^2)$, another 100 from $\mathcal{N}(1, \sigma^2)$. We also generate 100 anti-symmetric matrices $\gamma_{ij}$ from $\mathcal{N}(1, \sigma^2)$, that we denote in figure 3 $\mathcal{N}^{A}(1, \sigma^2)$. For each realization of $\gamma_{ij}$, we generate the transfer matrix and proceed as explained in section (3) to obtain the IEP in each case.

Figure 3 shows that for fitted data with a maximum entropy 1-time step Markov model, the IEP is zero only when all the measured 1-step correlations between neurons are symmetric, which is very unlikely for an experimental spike train. The degree of symmetry in the matrix of $\gamma$’s play an important role in the IEP.

5.4. Fourth example: Kinetic Ising model with random asymmetric interactions

This model of spike generation is an example of a non-equilibrium system, which has been used in [29] to approach the question of recovering the interactions of an asymmetrically-coupled Kinetic Ising model, with a time-independent external field to ensure stationarity. This is a discrete-time, synchronously updated Markov model in $\Sigma_N^1$, with transition matrix is given by:

$$P[\sigma' | \sigma] = \prod_{i=1}^{N} \frac{\exp[(2\sigma' - 1)\theta_i(\sigma)]}{2 \cosh[\theta_i(\sigma)]}, \forall \sigma, \sigma' \in \Sigma_N^1.$$  (38)
The fields \( h_i \) and the couplings \( J_{ij} \) are independent Gaussian variables and \( \alpha, \beta \in \mathbb{R} \). These set of stationary transition probabilities characterize an ergodic Markov chain with a unique invariant measure. Therefore, the scene is set to compute information entropy production under different scenarios.

In figure 4, for the Kinetic Ising model with random asymmetric interactions we recover the same structure found in figure 1 for the Integrate and Fire model, that is, the synaptic couplings are playing a major role in IEP, while the intensity of the stimulus is less relevant.

6. Discussion

One of the consequences of including non-synchronous constraints in the framework of the maximum entropy principle as constraints is that opens the possibility to broke the time-reversal symmetry introduced by time-independent models and thus capture the irreversible character of the underlying biological process, which is compulsory for a deeper understanding of the neural code and allows fit statistical models biologically more realistic. We have emphasized that the IEP is zero for time independent processes (time-reversible) derived from commonly used statistical models in this field, for example, Ising, \( K \)-pairwise, triplets, among others [5,28]. However, only time-dependent maximum entropy models induce time irreversible processes, feature highly expected from biological systems.

While many spiking neuronal network models consider the influence of pre-synaptic neurons, the most popular maximum entropy models in this field ignore them. Therefore, there is a clear phenomenological disagreement between these two different approaches, which as we show here and as mentioned before (see [30]) can be corrected including non-synchronous constraints.

When trying to explain the spike train statistics of networks of neurons using the maximum entropy principle, one hopes that the parameters of the potential shed light on the understanding of the nature of the spiking phenomena and the neural code, both clearly having a temporal structure. Perhaps the main message of this work is that limiting the complexity of the maximum entropy model using arguments of parsimony may be harmful when the underlying stochastic process generating the spikes is evidently time dependent.
18 of 20

**Figure 5.** IEP for the Kinetic Ising model with random asymmetric interactions. We consider $N = 6$. The components of field vector were drawn at random from a Gaussian $\mathcal{N}(-3, 1)$ and the coupling matrix $J_{ij}$ were drawn at random from a Gaussian $\mathcal{N}(0, 1)$. We plot the average value of IEP for 500 realizations of the synaptic coupling matrix for fixed $\alpha$ and $\beta$ in each case.

However, there are two main drawbacks of our approach. The first is inherited from the maximum entropy method that requires stationarity in the data. The second is that is based on the transfer matrix technique, so it may require an important computational effort for large-scale and long memory neuronal networks. There is a lot of room for progress going beyond the scope of this work, one possibility is to quantify the IEP for different choices of non-synchronous constraints and binning sizes on biological spike train recordings. A more ambitious goal would be to link the IEP as a signature of an underlying physiological process depending on time such as adaptation or learning. IEP is a much broader concept which can also be measured along non-stationary trajectories, thus IEP can be measured for time-dependent models where transition probabilities are explicitly given or can be computed (for example the Generalized Linear model [31]). Previous studies in the context of spike train statistics have measured the dynamical entropy production in spiking neuron networks using a deterministic approach based on the Pesin identity (sum of positive Lyapunov exponents) [32]. There are relationships between the deterministic and stochastic dynamics [33], and some interpretations of deterministic dynamical entropy production with information loss which should be investigated in more detail, in particular, if these relationships bring new knowledge in the field of computational neuroscience.

We have focused on spike train statistics, but our results are not restricted to this field and can be applied wherever Markov maximum entropy measures under constraints have to be inferred from data, especially for irreversible Markov chains from stochastic network theory [34], information theory [20], finance [21], among other disciplines.

**Acknowledgments:** We thank J.P Eckmann and Fernando Rosas for discussions and careful reading of the manuscript. RC was supported by an ERC advanced grant “Bridges”, CONICYT-PAI Insercion # 79160120. CM was supported by the CONICYT-FONDECYT Postdoctoral Grant No. 3140572.

**Author Contributions:** Both authors conceived the algorithm and wrote the and revised manuscript. All authors have read and approved the final manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.
Abbreviations

The following abbreviations are used in this manuscript:

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDPI</td>
<td>Multidisciplinary Digital Publishing Institute</td>
</tr>
<tr>
<td>DOAJ</td>
<td>Directory of open access journals</td>
</tr>
<tr>
<td>TLA</td>
<td>Three letter acronym</td>
</tr>
<tr>
<td>LD</td>
<td>Linear dichroism</td>
</tr>
</tbody>
</table>


