Article

Information Entropy Production of Maximum Entropy Markov Chains from Spike Trains

Rodrigo Cofre [1] and Cesar Maldonado [2]

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

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

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

1. Introduction

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

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

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

---

[1]   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] [2]. 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)[3][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] [4]. This

---

[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[5]. We denote the *spiking-state* of each neuron $\sigma_k^n = 1$ whenever the k-th neuron emits a spike at time n, and $\sigma_k^n = 0$ otherwise. The spike-state of the entire network at time n is denoted by $\sigma^n := \left[\sigma_k^n \right]_{k=1}^N$, which we call a *spiking pattern*. For $n_1 \leq n_2$, we denote by σ^{n_1, n_2} to an ordered concatenation of spike patterns

$$\sigma^{n_1, n_2} = \sigma^{n_1} \sigma^{n_1+1} \ldots \sigma^{n_2-1} \sigma^{n_2},$$

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 $\bar{\sigma}$. We denote the set of infinite binary sequences of N neurons Σ_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*:

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

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 ℓ.

---

[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 Σ_N^L be the state space of a discrete time Markov chain, and let us for the moment use the following notation $\sigma_{(n)} := \sigma^{n,n+L-1}$, for the random blocks and analogously $\omega_{(n)} := \omega^{n,n+L-1}$ 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,

$$\mathbb{P}\big[\sigma_{(n)} = \omega_{(n)} \mid \sigma_{(n-1)} = \omega_{(n-1)}, \ldots, \sigma_{(0)} = \omega_{(0)}\big] = \mathbb{P}\big[\sigma_{(n)} = \omega_{(n)} \mid \sigma_{(n-1)} = \omega_{(n-1)}\big]. \qquad (2)$$

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 $\sigma_{(0)} \to \tilde{\sigma}_{(1)}$ is *allowed* if they have the common sub-block $\sigma^{1,L-1} = \tilde{\sigma}^{1,L-1}$.

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

$$P_{\sigma_{(0)},\tilde{\sigma}_{(1)}} := \begin{cases} \mathbb{P}[\tilde{\sigma}_{(1)} \mid \sigma_{(0)}] > 0 & \text{if } \sigma_{(0)} \to \tilde{\sigma}_{(1)} \text{ is allowed} \\ 0, & \text{otherwise.} \end{cases} \qquad (3)$$

Note that P has $2^{NL} \times 2^{NL}$ entries, but it is a sparse matrix since each row has, at most, 2^N non-zero entries. Observe that by construction, for any pair of states there is a path of maximum length L in the graph of transition probabilities going from one state to the other, therefore the Markov chain is irreducible.

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 ν, the following Markov measure $\mu(\nu, P)$

$$\mu[\sigma_{(0)} = \omega_{(0)}, \sigma_{(1)} = \omega_{(1)}, \ldots, \sigma_{(k)} = \omega_{(k)}] = \nu(\omega_{(0)})P_{\omega_{(0)},\omega_{(1)}} \cdots P_{\omega_{(k-1)},\omega_{(k)}}, \qquad (4)$$

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 π such that

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

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 ν^n 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,

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

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

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

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 π is said to satisfy detailed balance if:

$$\pi_\ell P_{\ell,\ell'} = \pi_{\ell'} P_{\ell',\ell} \quad \forall \sigma^{(\ell)}, \sigma^{(\ell')} \in \Sigma_N^L. \qquad (7)$$

[121] If the detailed balance equations are satisfied, then the quantity inside the parenthesis in the right-hand
[122] side of (6) is zero.

3.4. Information Entropy rate and Information Entropy Production

[124] A well established measure of the amount of uncertainty of a probability measure ν is the
[125] *information entropy rate*, which we denote by $\mathcal{S}(\nu)$. In the case of independent sequences of spike
[126] patterns ($L = 1$), the entropy rate is given by:

$$\mathcal{S}(\nu) = -\sum_{\sigma^{(\ell)} \in \Sigma_N^1} \nu\left[\sigma^{(\ell)}\right] \log \nu\left[\sigma^{(\ell)}\right]. \tag{8}$$

[127]

[128]

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

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

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

[130]

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 ν in the
state space Σ_N^L at time n be given by

$$\mathcal{S}_n(\nu) = -\sum_{\sigma^{(\ell)} \in \Sigma_N^L} \nu^n(\sigma^{(\ell)}) \log \nu^n(\sigma^{(\ell)}).$$

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

$$\Delta\mathcal{S}_n := \mathcal{S}_{n+1}(\nu) - \mathcal{S}_n(\nu) = -\sum_{\sigma^{(\ell)} \in \Sigma_N^L} \nu^{n+1}(\sigma^{(\ell)}) \log \nu^{n+1}(\sigma^{(\ell)}) + \sum_{\sigma^{(\ell)} \in \Sigma_N^L} \nu^n(\sigma^{(\ell)}) \log \nu^n(\sigma^{(\ell)}).$$

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

$$\begin{aligned}\Delta\mathcal{S}_n = &-\sum_{\sigma^{(\ell)}, \sigma^{(\ell')} \in \Sigma_N^L} \nu^n(\sigma^{(\ell')}) P_{\ell', \ell} \log \frac{\nu^{n+1}(\sigma^{(\ell')}) P_{\ell', \ell}}{\nu^n(\sigma^{(\ell)}) P_{\ell, \ell'}} + \\ &\frac{1}{2} \sum_{\sigma^{(\ell)}, \sigma^{(\ell')} \in \Sigma_N^L} \left[\nu^n(\sigma^{(\ell')}) P_{\ell', \ell} - \nu^n(\sigma^{(\ell)}) P_{\ell, \ell'}\right] \log \frac{\nu^n(\sigma^{(\ell')}) P_{\ell', \ell}}{\nu^n(\sigma^{(\ell)}) P_{\ell, \ell'}},\end{aligned} \tag{10}$$

[131] the first part on the r.h.s of this equation is called *information entropy flow* and the second *information*
[132] *entropy production* [15].
[133] Observe that in the stationary state, one has that $\nu^n = \nu^{n+1} = \pi$, thus the change of entropy rate
[134] is zero, meaning that information entropy flow equal information entropy production, therefore is
[135] possible to attain a steady state of fixed maximum entropy, but having positive IEP. In this case we
[136] 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:

$$IEP(P, \pi) = \frac{1}{2} \sum_{\sigma^{(\ell)}, \sigma^{(\ell')} \in \Sigma_N^L} \left[\pi_{\ell'} P_{\ell', \ell} - \pi_\ell P_{\ell, \ell'}\right] \log \frac{\pi_{\ell'} P_{\ell', \ell}}{\pi_\ell P_{\ell, \ell'}} \geq 0, \tag{11}$$

137 nevertheless, we stress the fact that one can obtain the information entropy production rate also in the
138 non-stationary case.

3.5. Example: Discrete time spiking neuronal network model

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

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

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

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

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

$$V(t+1) = F(V(t)) + \sigma_B B(t) \tag{12}$$

156 where

$$F_i(V(t)) = \gamma V_i(t)\big(1 - Z[V_i(t)]\big) + \alpha \sum_{j=1}^{N} W_{ij} Z[V_j(t)] + \beta I_i. \tag{13}$$

157 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
158 $Z[V_i(t)] = 1$ one says that neuron i *spike* otherwise is *silent*. We extend the definition of Z to vectors:
159 $Z[V(t)]$ is the vector with components $Z[V_i(t)], i = 1, ..., N$. The *leak rate* is denoted by $\gamma \in [0, 1[$,
160 and W_{ij} is called the *synaptic weight* from the neuron j to the neuron i. The synaptic weight is said to
161 be *excitatory* if $W_{ij} > 0$ or *inhibitory* if $W_{ij} < 0$. The components of the vector $B(t)$ are independent
162 normalized Gaussian random variables and σ_B is the noise amplitude parameter. The parameters
163 α and β were introduced in order to control the intensity of the synaptic weights and the stimulus,
164 respectively.

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

$$P[\sigma \mid \sigma'] = \prod_{i=1}^{N} \left[\sigma_i \, \varphi\Big(\frac{\theta - C_i(\alpha, \beta, \sigma')}{\sigma_B}\Big) + (1 - \sigma_i)\Big(1 - \varphi\Big(\frac{\theta - C_i(\alpha, \beta, \sigma')}{\sigma_B}\Big)\Big) \right], \tag{14}$$

where,

$$C_i(\alpha, \beta, \sigma') = \gamma \, \alpha \sum_{j=1}^{N} W_{ij} \sigma'_j + \beta I_i \tag{15}$$

and

$$\varphi(x) = \int_x^\infty e^{\frac{-u^2}{2}} \, du. \tag{16}$$

170 The function C takes into account the past and the external stimuli (see [17] for details). These
171 conditional (transition) probabilities define an ergodic Markov chain, specified by the biophysical
172 dynamics of the spiking network. From the transition probabilities (14) and its unique steady state, we
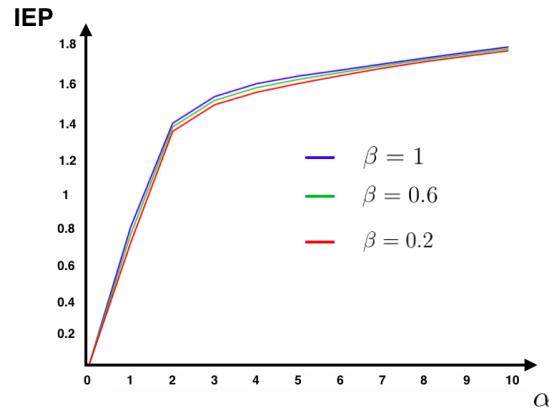173 compute the IEP of this model using (11) for different values of the parameters α and β (see Figure 1).

Figure 1. Plot of the average value of IEP for 500 realizations of the synaptic weight matrix for fixed α and β in each case. We fix the following values of the parameters: $N = 6, \gamma = 0.2, \sigma_b = 1, \theta = 1, I_i = 1 \quad \forall i \in \{1,..,6\}$. 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 α and β in each case..

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

184

185 **4. Theoretical Framework**

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

196 *4.1. Inference of the maximum entropy Markov process*

197 The problem of estimating the Markov chain of maximum entropy constrained by the data is of
198 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 Σ_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-translation invariant observables*), i.e., for any time-step n, $f(\sigma^{0,R-1}) = f(\sigma^{n,n+R-1})$ whenever $\sigma^{0,R-1} = \sigma^{n,n+R-1}$. Examples of observables are products of the form:

$$f(\sigma^{0,T}) = \prod_{u=1}^{r} \sigma_{k_u}^{n_u}, \tag{17}$$

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:

$$\mathcal{H}(\sigma^{(\ell)}) := \sum_{l=1}^{2^{NR}} h_l m_l(\sigma^{(\ell)}) \quad \sigma^{(\ell)} \in \Sigma_N^R, \tag{18}$$

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

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

4.2.1. Additive observables of spike trains

Let ϕ 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 $\{f \circ \phi^i(\sigma)\}$ as a random variable whose statistical properties depend on those of the process producing the samples of σ 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 σ, and let n be the sample length. The average of the observable f of range $R \geq 1$ in σ 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.

$$A_n(f) = \frac{1}{n-R+1} \sum_{i=0}^{n-R} f \circ \phi^i(\bar{\sigma}),$$

229 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

231   Let $A_n(f_k) = C_k$ be the average value of K observables for $k \in \{1, \ldots, K\}$. As the empirical
232 average of monomials is not enough to uniquely determine the spike train statistics (there are infinitely
233 many probability measures sharing the same averages of monomials), we use the maximum entropy
234 method to obtain the Markov measure μ that maximizes the KSE among all measures ν that match the
235 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
236 following variational problem under constraints:

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

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

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

243 where $\mathcal{P}[\mathcal{H}]$ is called the *free energy or topological pressure*, $\mathcal{M}_{inv}$ is the set of invariant measures with
244 respect to the shift ϕ and $\nu[\mathcal{H}] = \sum_{k=1}^{K} h_k \nu[f_k]$ is the average value of $\mathcal{H}$ with respect to ν.
245   In this paper, we only consider potentials $\mathcal{H}$ of finite range, for which there is a unique measure μ
246 attaining the supremum [22] and is a **Gibbs measure in the sense of Bowen.**

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

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

252 It is easy to see that the classical form of Boltzmann-Gibbs distributions $\mu[\sigma] = e^{\mathcal{H}(\sigma)}/Z$ is a particular
253 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

255   The functional $\mathcal{P}[\mathcal{H}]$ has the following property:

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

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

260 (21). Efficient algorithms exist to estimate the Lagrange multipliers for the maximum entropy problem
261 with non-synchronous constraints [10].

262

4.4. Ruelle-Perron-Frobenius transfer operator

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

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

266

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

$$\mathcal{L}_{\mathcal{H}}(\ell, \ell') = \begin{cases} e^{\mathcal{H}(\sigma^{0,L})} & \text{if } \sigma^{0,L} \sim \sigma^{(\ell)} \to \sigma^{(\ell')} \\ 0, & \text{otherwise.} \end{cases} \tag{24}$$

269

270 From the assumption $\mathcal{H} > -\infty$, each allowed transition corresponds to a positive entry in the matrix
271 $\mathcal{L}_{\mathcal{H}}$.

4.5. Maximum entropy Markov chain for finite range potentials

273   The matrix (24) is primitive[8] by construction, thus it satisfies the Perron-Frobenius theorem [25].
274 Let $\rho > 0$ be its spectral radius. Because of the irreducibility of the transfer matrix, ρ is an eigenvalue
275 of multiplicity 1 strictly larger in modulus than the other eigenvalues. For every $\sigma^{(\ell)} \in \Sigma_N^L$, let us
276 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
277 eigenvalue ρ. Notice that $L_\ell > 0$ and $R_\ell > 0$ for all $\sigma^{(\ell)} \in \Sigma_N^L$. Using spectral properties of the transfer
278 matrix, we obtain the maximum entropy Markov transition probability matrix [22]:

$$P_{\ell,\ell'} := \frac{\mathcal{L}_{\mathcal{H}}(\ell, \ell') R_{\ell'}}{R_\ell \rho}, \quad \forall \sigma^{(\ell)}, \sigma^{(\ell')} \in \Sigma_N^L. \tag{25}$$

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

$$\pi_\ell := \frac{L_\ell R_\ell}{\langle L, R \rangle}, \quad \forall \sigma^{(\ell)} \in \Sigma_N^L. \tag{26}$$

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

4.6. IEP of the inferred Markov maximum entropy process

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

$$IEP(\mathcal{L}_{\mathcal{H}}) = \sum_{\sigma^{(\ell)}, \sigma^{(\ell')} \in \Sigma_N^L} \frac{L_\ell}{\langle L, R \rangle} \frac{\mathcal{L}_{\mathcal{H}}(\ell, \ell') R_{\ell'}}{\rho} \log \left[\frac{L_\ell R_{\ell'} \mathcal{L}_{\mathcal{H}}(\ell, \ell')}{L_{\ell'} R_\ell \mathcal{L}_{\mathcal{H}}(\ell', \ell)} \right] \tag{27}$$

---

8   The matrix A is primitive if there is an $n \in \mathbb{N}$, s.t. A^n has only positive components

286   This is a quantity of major interest in spike train statistics, as it measure the degree of time
287 irreversibility of the inferred maximum entropy Markov chain. Although it is a straightforward result,
288 it is quite general and of practical use, as we will see in the examples below. We can apply (25) and (26)
289 to equation (7), we obtain the detailed balance condition in terms of the transfer matrix and its spectral
290 properties:

$$\frac{L_\ell R_\ell}{\langle L, R \rangle} \frac{\mathcal{L}_\mathcal{H}(\ell, \ell') R_{\ell'}}{R_\ell \, s} = \frac{L_{\ell'} R_{\ell'}}{\langle L, R \rangle} \frac{\mathcal{L}_\mathcal{H}(\ell', \ell) R_\ell}{R_{\ell'} \, s}$$

291

292 Simplifying we obtain:

$$\frac{\mathcal{L}_\mathcal{H}(\ell, \ell')}{\mathcal{L}_\mathcal{H}(\ell', \ell)} = \frac{R_\ell \, L_{\ell'}}{R_{\ell'} \, L_\ell} \tag{28}$$

293 *4.7. Large deviations for observables of maximum entropy Markov chains*

294   The goal of large deviations is to compute the asymptotic probability distribution $\mathbb{P}(A_n(f) = s)$
295 for a given finite range observable f and for $s \neq \mathbb{E}(f)$. More precisely, we say that $\mathbb{P}(A_n(f))$ satisfies a
296 large deviation principle with rate $I_f(s)$ if the following limit exists,

$$\lim_{n \to \infty} -\frac{1}{n} \ln \mathbb{P}(A_n(f) = s) = I_f(s).$$

297 In which the dominant behavior of $\mathbb{P}(A_n(f))$ is decaying exponentially fast with the sample size n, as

$$\mathbb{P}(A_n(f) = s) \approx e^{-n I_f(s)}. \tag{29}$$

298 We define the *scaled cummulant generating function* (SCGF) associated to the random variable
299 (observable) f denoted by $\lambda_f(k)$ as follows,

$$\lambda_f(k) := \lim_{n \to \infty} \frac{1}{n} \ln \mathbb{E}\left[e^{n k A_n(f)} \right], \quad k \in \mathbb{R}. \tag{30}$$

300 The n-th cumulant of the random variable f can be obtained by differentiating $\lambda_f(k)$ with respect to
301 k, n times and evaluating the result at $k = 0$. The next theorem by Gärtner-Ellis theorem relates the
302 SCGF and the large deviations rate function. The Gärtner-Ellis theorem relies on the differentiability of
303 $\lambda_f(k)$, which is guaranteed for finite state Markov chains [26]. This theorem has several formulations,
304 which usually require some technical definitions beforehand. Here we stated it in a simplified form,
305 which is what we need for our purposes.

306

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

$$I_f(s) = \max_{k \in \mathbb{R}} \{ ks - \lambda_f(k) \} \tag{31}$$

309 The Gärtner-Ellis Theorem is very useful in our context, because it bypasses the direct calculation of
310 $\mathbb{P}(A_n(f))$ in (29), i.e., having $\lambda_f(k)$ a simple calculation leads to the rate function of f. As we will see
311 in the next section $\lambda_f(k)$ naturally appears in the context of Maximum entropy Markov chains.

312 *4.8. Large deviations for the IEP*

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

$$\tilde{P}^{(f)}_{\ell, \ell'}(k) = P_{\ell, \ell'} e^{k f(\ell')} \tag{32}$$

315 or for a two time step observable:

$$\tilde{P}^{(f)}_{\ell,\ell'}(k) = P_{\ell,\ell'}e^{kf(\ell,\ell')} \tag{33}$$

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

$$\tilde{P}^{(f)}_{\ell,\ell'}(k) = \frac{\mathcal{L}_{\mathcal{H}}(\ell,\ell')R_{\ell'}}{R_\ell\,\rho}e^{kf(\ell,\ell')} \tag{34}$$

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

$$\lambda_f(k) = \ln(\rho(\tilde{P}^{(f)})) \tag{35}$$

321 This result is valid if the state-space of the Markov chain is finite, where it can be proved furthermore
322 that $\lambda_f(k)$ is differentiable and $\lambda'_f(0) = \mathbb{E}(f)$.
323
324 **Remark:** The observable f does not need to belong in the set $\{f_k\}_{k=1}^K$ of chosen observables to fit the
325 Markov maximum entropy process. We denote $\rho(\tilde{P}^{(f)})$ the dominant eigenvalue (i.e., with largest
326 magnitude) of the matrix $\tilde{P}^{(f)}$, which is unique by the Perron-Frobenius theorem.
327
328   We are interested in the fluctuations of the IEP. For that purpose we define the following
329 observable:

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

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

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

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

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

334 where,

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

335 which is a matrix of positive elements.
336
337   Using the Gärtner-Ellis theorem we obtain the rate function $I_W(s)$ for the IEP observable:

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

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

$$\lambda_W(k) = \lambda_W(-k-1)$$

339

340 Since $\lambda'_W(0) = IEP(\pi, P)$ the symmetry implies

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

341 4.8.1. Gallavotti-Cohen fluctuation theorem

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

$$\frac{P\left[\frac{W_n}{n} \approx s\right]}{P\left[\frac{W_n}{n} \approx -s\right]} \asymp e^{ns}$$

345 This means that the positive fluctuations of $\frac{W_n}{n}$ are exponentially more probable than negative
346 fluctuations of equal magnitude. This is a universal ratio, i.e., no free parameters are involved
347 and is experimentally observable.

5. Examples

348

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

353 *5.1. First example: Toy model*

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

$$\mathcal{H}(\sigma^{0,1}) = h_1 \sigma_1^1 \sigma_2^0.$$

The transfer matrix (24) associated to $\mathcal{H}$ is in this case a 4×4 matrix:

$$\mathcal{L}_{\mathcal{H}} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & e^{h_1} & 1 & e^{h_1} \\ 1 & e^{h_1} & 1 & e^{h_1} \end{pmatrix}.$$

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

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

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

From the spectral properties of $\mathcal{L}_{\mathcal{H}}$ we obtain the Markov transition matrix (25), which reads,

$$P_{\sigma^0, \sigma^1} = \frac{1}{\rho} \begin{pmatrix} 1 & 1 & \frac{1+e^{h_1}}{2} & \frac{1+e^{h_1}}{2} \\ 1 & 1 & \frac{1+e^{h_1}}{2} & \frac{1+e^{h_1}}{2} \\ \frac{2}{1+e^{h_1}} & \frac{2e^{h_1}}{1+e^{h_1}} & 1 & e^{h_1} \\ \frac{2}{1+e^{h_1}} & \frac{2e^{h_1}}{1+e^{h_1}} & 1 & e^{h_1} \end{pmatrix},$$

The unique invariant measure of this irreducible Markov transition matrix is given by equation (26), and its entries are given by,

$$\pi\left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right) = \frac{4}{\rho^2}, \quad \pi\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) = \frac{2(\rho-2)}{\rho^2}, \quad \pi\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) = \frac{2(\rho-2)}{\rho^2}, \quad \pi\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right) = \frac{(\rho-2)^2}{\rho^2}.$$

358   It is easy to check that π is invariant w.r.t. the transition matrix P, that is $\pi P = \pi$.

From this example, we can verify that *generically the detailed balance condition is not satisfied*; for example:

$$P\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix} \Big| \begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)\pi\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) \neq P\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix} \Big| \begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)\pi\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right).$$

359   As we can see in figure 2, the maximum entropy measure for the unconstrained problem is attained at
360   the uniform distribution as expected ($h_1 = 0$, eigenvalue $\rho = 4$ assigning probability $\frac{1}{4}$ to each spike
361   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 $f = \sigma_1^1 \sigma_2^0$,

$$A_T(f) = 0.1$$

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

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

362   Solving we find $h_1 = -1.09861$. Therefore, among all the Markov chains that match exactly the
363   restriction, the one that maximizes the information entropy is the one obtained by fixing h_1 at the
364   found value. Is easy to check that the variational principle (21) is satisfied.
365   From the transition probability matrix P and the invariant measure π, we compute the KSE (9)
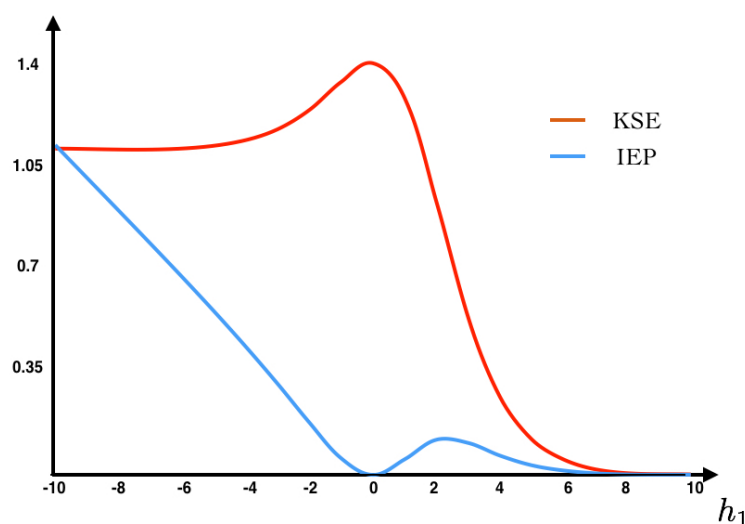366   and the IEP (27) as a function of the parameter h_1 (see figure 2).

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 $h_1 = 0$.
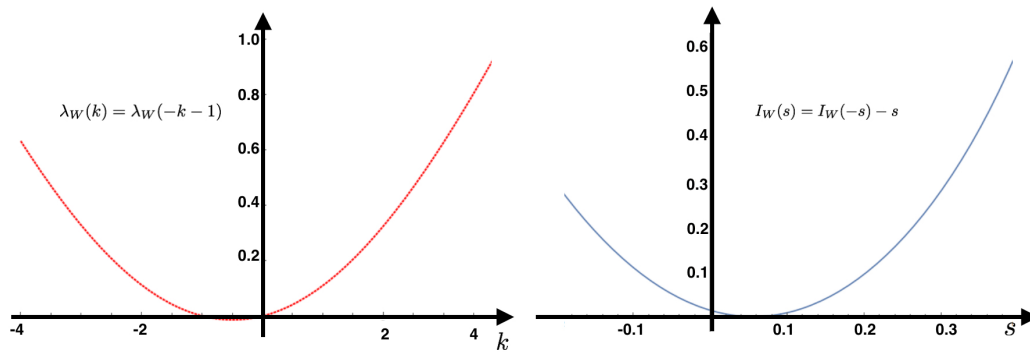
Figure 3. Gallavotti-Cohen fluctuation theorem for the first example at the parameter value $h_1 = -1$. Left) We show the SCGF associated to W, $\lambda_W(k)$, the derivative at zero is the IEP of the maximum entropy Markov chain, which in this case is 0.0557. This value coincides with the minimum of 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 $\sigma \to \sigma'$; $\sigma, \sigma' \in \Sigma_N^1$ are considered and all transitions are allowed. In this case, the potential does not "see" the past i.e. $\mathcal{L}_{\mathcal{H}}(\sigma, \sigma') = e^{\mathcal{H}(\sigma')}$. The matrix $\mathcal{L}_{\mathcal{H}}$ has a unique maximum eigenvalue:

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

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

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

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

$$P(\sigma' \mid \sigma) = P(\sigma') = \pi(\sigma') = \frac{e^{\mathcal{H}(\sigma')}}{Z}; \quad \forall \sigma, \sigma' \in \Sigma_N^1, \tag{36}$$

In this case, the invariant measure π 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:

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

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 $\mathcal{D}^O$. Let $g : \{0, \dots, T\} \to \{0, \dots, T\}$ be a function that randomly shuffles the time indexes, we call $\mathcal{D}^{RS}$ the data set obtained after this transformation. Finally consider $\mathcal{D}^I$, the data set with inverted time indexes,

$$\mathcal{D}^O = \{\sigma^0, \sigma^1, \sigma^2, \dots, \sigma^{T-1}, \sigma^T\}$$
$$\mathcal{D}^{RS} = \{\sigma^{g(0)}, \sigma^{g(1)}, \sigma^{g(2)}, \dots, \sigma^{g(T-1)}, \sigma^{g(T)}\}$$
$$\mathcal{D}^I = \{\sigma^T, \sigma^{T-1}, \sigma^{T-2}, \dots, \sigma^1, \sigma^0\}.$$

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:

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

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 γ_{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 γ_{ij} is generated, if γ_{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 $\gamma_{ij} = 0, \forall i, j \in \{1, \dots, N\}$, where we recover the Ising model. Next, we fix the values of h_i and J_{ij} (random values), and we generate 100 matrices γ_{ij} by drawing their components from Gaussian distributions $\mathcal{N}(0, e^2)$, another 100 from $\mathcal{N}(1, e^2)$. We also generate 100 anti-symmetric matrices γ_{ij} from $\mathcal{N}(1, e^2)$, that we denote in figure 3 $\mathcal{N}^A(1, e^2)$. For each realization of γ_{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 γ'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 Σ_N^1 with transition matrix is given by:

$$P[\sigma' \mid \sigma] = \prod_{i=1}^{N} \frac{\exp[(2\sigma_i' - 1)\theta_i(\sigma)]}{2\cosh[\theta_i(\sigma)]}, \quad \forall \sigma, \sigma' \in \Sigma_N^1 \tag{38}$$
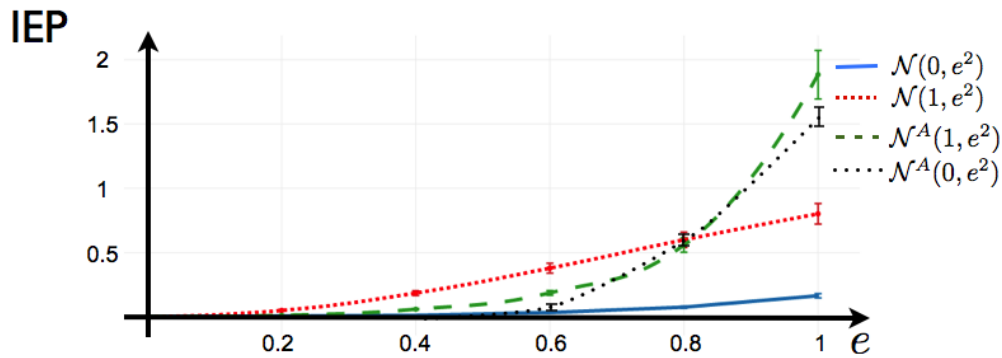
Figure 4. IEP for the 1-time step Markov potential. The parameters h_i and J_{ij} are draw at random one time and remain fixed. We draw at random the components of 100 matrices γ_{ij} from a Gaussian distribution with different values of mean and standard deviation e. We plot the average value of IEP for each case, with the respective error bars.

$$\theta_i(\sigma) = \beta\, h_i + \alpha \sum_{j=1}^{N} J_{ij}(2\sigma_i - 1) \quad \forall \sigma \in \Sigma_N^1. \tag{39}$$

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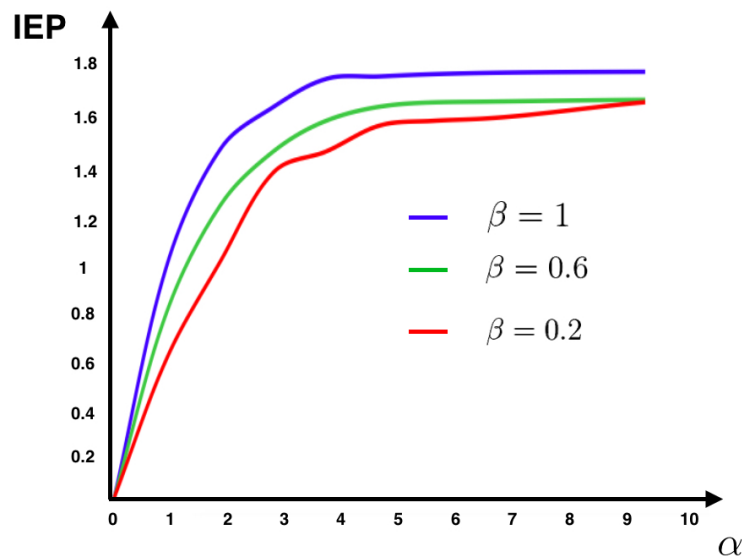
Figure 5. IEP for the Kinetic Ising model with random asymmetric interactions. We consider $N = 6$. The components of field vector were drawn at random from a Gaussian $\mathcal{N}(-3, 1)$ and the coupling matrix J_{ij} were drawn at random from a Gaussian $\mathcal{N}(0, 1)$. We plot the average value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

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:

MDPI	Multidisciplinary Digital Publishing Institute
DOAJ	Directory of open access journals
TLA	Three letter acronym
LD	linear dichroism

1. Palsso, B. *Systems Biology: Properties of Reconstructed Networks*; Cambridge University Press, 2006.
2. Deem, M. Mathematical adventures in biology. *Phys Today* **2007**, *60*, 42–47.
3. Shi, P.; Qian, H., Frontiers in Computational and Systems Biology, J. Feng, W. Fu and F. Sun Eds; Springer, 2010; chapter Irreversible Stochastic Processes, Coupled Diffusions and Systems Biochemistry., pp. 175–201.
4. Jiang, D.Q.; Qian, M.; Qian, M.P. *Mathematical Theory of Nonequilibrium Steady States*; Springer, 2004.
5. Schneidman, E.; Berry II, M.; Segev, R.; Bialek, W. Weak pairwise correlations imply string correlated network states in a neural population. *Nature* **2006**, *440*, 1007– 1012.
6. Pillow, J.W.; Shlens, J.; Paninski, L.; Sher, A.; Litke, A.M.; Chichilnisky, E.J.; Simoncelli, E.P. Spatio-temporal correlations and visual signaling in a complete neuronal population. *Nature* **2008**, *454*, 995–999.
7. Tkačik, G.; O.Marre.; D.Amodei.; E.Schneidman.; W.; 2nd, M.B. Searching for collective behavior in a large network of sensory neurons. *Plos Computational Biology* **2013**, *10*.
8. Vasquez, J.; Marre, O.; Palacios, A.; Berry II, M.; Cessac, B. Gibbs distribution analysis of temporal correlation structure on multicell spike trains from retina ganglion cells. *J. Physiol. Paris* **2012**. in press.
9. Ganmor, E.; Segev, R.; Schneidman, E. The architecture of functional interaction networks in the retina. *The journal of neuroscience* **2011**, *31*, 3044–3054.
10. Nasser, H.; Cessac, B. Parameter Estimation for Spatio-Temporal Maximum Entropy Distributions: Application to Neural Spike Trains. *Entropy* **2014**, *16*.
11. Bowen, R. *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, revised ed.; Vol. 470, *Lecture Notes in Mathematics*.
12. Fernandez, R.; Maillard, G. Chains with complete connections : General theory, uniqueness, loss of memory and mixing properties. *J. Stat. Phys.* **2005**, *118*, 555–588.
13. Cessac, B.; Cofré, R. Spike train statistics and Gibbs distributions. *J. Physiol. Paris* **2013**, *107*, 368–368.
14. Galves, A.; Löcherbach, E. Infinite Systems of Interacting Chains with Memory of Variable Length-A Stochastic Model for Biological Neural Nets. *Journal of Statistical Physics* **2013**, *151*, 896–921.
15. Gaspard, P. Time-reversed dynamical entropy and irreversibility in Markovian random processes. *J. Statist. Phys.* **2004**, *117*, 599–615.
16. Kitchens, B.P. *Symbolic Dynamics: One-sided, Two-sided and Countable State Markov Shifts*; Springer-Verlag, 1998.
17. Cessac, B. A discrete time neural network model with spiking neurons. Rigorous results on the spontaneous dynamics. *J. Math. Biol.* **2008**, *56*, 311–345.
18. Ruelle, D. *Thermodynamic formalism*; Addison-Wesley,Reading, Massachusetts, 1978.
19. Jaynes, E. Information theory and statistical mechanics. *Phys. Rev.* **1957**, *106*.
20. der Straeten, E.V. Maximum Entropy Estimation of Transition Probabilities of Reversible Markov Chains. *Entropy* **2009**, *4*.
21. Chliamovitch, G.; Dupuis, A.; Chopard, B. Maximum Entropy Rate Reconstruction of Markov Dynamics. *Entropy* **2015**, *6*.
22. Baladi, V. *Positive Transfer Operators and Decay of Correlations*; Vol. 16, World Scientific, 2000.
23. Chazottes, J.R.; Ramirez, L.; Ugalde, E. Finite type approximations of Gibbs measures on sofic subshifts. *Nonlinearity* **2005**, *18*, 445–463.
24. Maldonado, C.; Salgado-García, R. Markov approximations of Gibbs measures for long-range interactions on 1D lattices. *J. Stat. Mech. Theory Exp.* **2013**, pp. P08012, 14.
25. Gantmacher, F.R. *The theory of matrices*; AMS Chelsea Publishing, Providence, RI, 1998.
26. Lancaster, P. *Theory of matrices*; Academic Press, 1969.

27. Marre, O.; El Boustani, S.; Frégnac, Y.; Destexhe, A. Prediction of spatiotemporal patterns of neural activity from pairwise correlations. *Physical review letters* **2009**, *102*.

28. Tkačik, G.; Mora, T.; O.Marre.; D.Amodei.; Palmer, S.; II, M.B.; Bialek, W. Thermodynamics and signatures of criticality in a network of neurons. *PNAS* **2015**, *112*.

29. Roudi, Y.; Hertz, J. Mean Field Theory For Non-Equilibrium Network Reconstruction. *Phys. Rev. Lett.* **2011**, *106*.

30. Cofré, R.; Cessac, B. Exact computation of the maximum entropy potential of spiking neural networks models. *Physical Review E* **2014**, *107*, 368–368.

31. Pillow, J.; Paninski, L.; Uzzell, V.; Simoncelli, E.; Chichilnisky, E. Prediction and decoding of retinal ganglion cell responses with a probabilistic spiking model. *J. Neurosci* **2005**, *25*, 11003–11013.

32. Monteforte, M.; Wolf, F. Dynamical Entropy Production in Spiking Neuron Networks in the Balanced State. *Physical Review Letters* **2010**, *10*.

33. Gaspard, P. Time asymmetry in nonequilibrium statistical mechanics. *Advances in Chemical Physics* **2007**, *135*, 83–133.

34. J. Delvenne, A.L. Centrality measures and thermodynamic formalism for complex networks. *Phys. Rev. E* **2011**.